Estimation of Premorbid Intelligence in Traumatic Brain Injury: NART and AUSNART Performance in an Australian Sample.

Tracey A. Dean BA (Hons)

Submitted in fulfillment of the requirements for the Degree of

Doctor of Psychology

School of Psychology

University of Tasmania

October 2009

This thesis contains no material, which has been accepted for a degree or diploma by the University or any other institution. To the best of my knowledge and belief this thesis contains no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Date: 20 · 11 · 2009

Tracey Dean

This thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968.*

Date: 20 · 11 · 2009

Tracey Dean

Acknowledgements

I would like to express my sincere thanks to so many people, without whom this research would not have been possible. First my supervisor, Clive Skilbeck, who has been an incredible source of knowledge and wisdom; who always remained positive, supportive and patient and was always available for my incessant queries.

I would like to thank all staff at the Neurotrauma Register Research (NTR), particularly Avril and Sama who's never failing support, encouragement and friendship has been a motivating force. To the patients who gave up their time for research, often in difficult circumstances, I would like to say thank you as without you this research would not be possible.

I would also like to thank staff and colleagues at the University of Tasmania, particularly Jan Martin and Matt Thomas who have always provided support and encouragement and made the journey an enjoyable one.

Finally I would like to sincerely thank my wonderful family and friends for their consistent unfailing words of encouragement and support for many years, particularly my parents, my sister and my daughters, Elle and Laura who have been beside me for this long journey. To my husband, David, I sincerely thank you for the endless love, support and encouragement and to Chris, Alex and Grace thank you also for your patience and kind words of encouragement.

Table of Contents

Abstract		1
Chapter 1	Overview of the Thesis	3
Chapter 2	Introduction to Traumatic Brain Injury	8
	2.1 Epidemiology	8
	2.2 Traumatic brain injury defined	9
	2.3 Types of brain trauma	10
	2.4 Measuring severity of injury	12
	2.5 Outcome following brain injury	15
	2.6 Summary	18
Chapter 3	Estimation of Premorbid Ability	20
	3.1 Demographic-predicted intelligence estimate	21
	3.2 'Hold'/'don't hold' methods	23
	3.3 Best performance method	24
	3.4 Combined demographic-current measures	26
	3.5 Reading ability	29
	3.5.1 Schonell Graded Word Reading Test	31
	3.5.2 Wide Range Achievement Test	32
	3.5.3 Cambridge Contextual Word Reading Test	34
	3.5.4 Wechsler Test of Adult Reading	35
	3.6 Lexical decision making task	38
	3.7 Summary	39
Chapter 4	Development of the National Adult Reading Test	41
	4.1 Reliability	46
	4.2 Validity	48
	4.3 NART & WAIS-R validation	52
	4.4 NART & WAIS-III validation	55
	4.5 Validation of NART with clinical groups	55
	4.6 Validity of the short-NART	59

	4.7	Combined NART-demographic regression equations	63
	4.8	NART adaptations	67
	4.9	The Australian National Adult Reading Test (AUSNART)	67
	4.10	NART limitations	70
	4.11	Summary	71
Chapter 5	Can NART performance be Impaired?		73
	5.1 Summary		79
Chapter 6	NART Performance in Traumatic Brain Injury		81
	6.1 Summary		89
Chapter 7	Study 1: Examining NART Sensitivity to TBI		92
	7.1 Aims and hypotheses		92
	7.2 Method		94
		7.2.1 Participants	94
		7.2.2 Materials	96
		7.2.3 Procedure	96
		7.2.4 Design and analyses	97
	7.3 Results		98
		7.3.1 Descriptive statistics	98
		7.3.2 Severity & age analyses	101
		7.3.3 Severity & education analyses	104
		7.3.4 Severity & socio-economic status (SES) analyses	111
		7.3.5 Severity & gender analyses	114
	7.4 Discussion		115
		7.4.1 Limitations	122
Chapter 8	Study 2: Examining the Australian NART (AUSNART) in a TBI Population: Effects of Demographic Variables and Comparison with NART		126
	8.1 Aims and hypotheses		129
	8.2 Method		130
		8.2.1 Participants	130

	8.2.2 Materials	131
	8.2.3 Procedure	131
	8.2.4 Design and analyses	132
	8.3 Results	133
	8.3.1 Descriptive statistics	133
	8.3.2 Severity & age analyses	136
	8.3.3 Severity & education analyses	137
	8.3.4 Severity & gender analyses	141
	8.3.5 Comparison of NART & AUSNART performance	142
	8.4 Discussion	144
	8.4.1 Limitations	148
Chapter 9	Study 3: Predicting NART Performance in Traumatic Brain Injury	150
	9.1 Aims and hypotheses	153
	9.2 Method	153
	9.2.1 Participants	153
	9.2.2 Materials	154
	9.2.3 Procedure	154
	9.2.4 Design and analyses	154
	9.3 Results	155
	9.3.1 Descriptive statistics	155
	9.3.2 Correlations	155
	9.3.3 Multiple Regression	156
	9.3.4 Accuracy rates	160
	9.3.5 Mean Change prediction method	162
	9.3.6 Accuracy rate for the mean change prediction method	164
	9.4 Discussion	166
	9.4.1 Limitations	172
Chapter 10	General Discussion	174
	10.1 Overview of the findings	177
	10.2 Clinical implications	187
	10.3 Future research	188

References			190
Appendix A			204
	A1	NART Word List	205
	A2	NART errors and estimated WAIS-III FSIQ scores	206
	A3	AUSNART Word List	207
	A4	AUSNART errors and estimated FSIQ scores	208
	A5	Conversion of obtained NART error score to predicted NART error score	210
Appendix B		Study1	CD
	B1	Raw Data	
	B2	Frequencies	
	B3	Severity & age SPSS print out	
	B4	Severity & education SPSS print out	
	B5	Severity & SES SPSS print out	
	B6	Severity & gender SPSS print out	
Appendix C		Study 2	CD
	C1a	Raw data for AUSNART analyses	
	C1b	Raw data for AUSNART & NART analyses	
	C2	AUSNART frequencies	
	C3	Severity & age SPSS print out	
	C4	Severity & education SPSS print out	
	C5	Severity & gender SPSS print out	
	C6	AUSNART & NART correlations	
	C7	AUSNART & NART frequencies	
	C8	AUSNART & NART repeated measures & t-tests	
Appendix D		Study 3	CD
	D1	Raw Data for Study 3	
	D2	Correlations	
	D3	Multiple Regression	
	D4	Multiple Regression Education Level 1	

D5 Multiple Regression Education Level 2
D6 Prediction equation accuracy rates
D7 Accuracy rates for FSIQ groups
D8 Mean Change Score Method analyses

List of Tables

Chapter 7 – Study 1

Table 7.1

Descriptive statistics for the total sample (N = 194).

Table 7.2

Frequencies and percentages of cause of injury.

Table 7.3

Number of participants in each age group for each severity group.

Table 7.4a

Means (standard deviations) of NART error scores according to age for all three assessments.

Table 7.4b

The t-values and significance levels of age group comparisons of NART errors for all three assessments.

Table 7.5

Number of participants in each education group for each severity group.

Table 7.6a

Means (standard deviations) of NART errors according to years of education and severity for all three assessments.

Table 7.6b

The t values and significance levels for differences in NART errors between education groups for each severity group for all three assessments.

Table 7.7a

The t values and significance levels for comparisons of NART errors according to severity for participants with ≤11 years education (n = 87) for all three assessments.

Table 7.7b

The t values and significance levels for comparisons of NART errors according to severity for participants with >11 years education (n = 107) for all three assessments.

Table 7.8

Number of participants in each SES group for each severity group for three assessments.

Table 7.9

The t-values and significance levels for SES group comparisons for all three assessments.

Chapter 8 – Study 2

Table 8.1

Descriptive statistic for the total AUSNART sample (N = 92).

Table 8.2

Frequencies and percentages of cause of injury for the total AUSNART sample (N = 92).

Table 8.3

Frequencies and percentages of demographic/clinical variables for the sample who completed both NART and AUSNART at each assessment (N = 88).

Table 8.4

Number of participants in each age group for each severity group.

Table 8.5

Number of participants in each education group for each severity group.

Table 8.6

The t-values and significance levels comparisons of AUSNART errors for each age group for all three assessments.

Table 8.7
Means, standard deviations and correlations (Pearson's) for NART and AUSNART estimated FSIQ.

Table 8.8
Mean (standard deviation) and t-values for comparisons of NART and AUSNART estimated FSIQ for each assessment.

Chapter 9 – Study 3

Table 9.1
Correlations between NART score and estimated FSIQ with education and age for the total sample (N = 194).

Table 9.2
Multiple Regression for predicting NART initial errors from demographic variables.

Table 9.3
Multiple Regression for predicting 12 month NART errors from demographic variables.

Table 9.4
Multiple Regression for predicting 12 month NART errors from the initial NART error scores and demographic variables.

Table 9.5
Percentages and frequencies (n) of predicted – obtained difference scores

Table 9.6
Percentages and frequencies (n) of predicted - obtained difference scores according to FSIQ for the NART initial error prediction equation.

Table 9.7
Means (standard error) for the initial impaired NART, 12 month recovered NART and mean NART improvement over 12 months for each education level and each age group.

Table 9.8

Means (standard error) for the initial estimated FSIQ, 12 month estimated FSIQ and mean IQ improvement over 12 months for each education level and each age group.

Table 9.9

Percentages and frequencies (n) of predicted – obtained difference for each education level and age group according to mean change prediction method.

List of Figures

Chapter 7 - Study 1

Figure 7.1 Mean NART errors according to years of education for each assessment.

Figure 7.2 Mean NART errors according to SES for each assessment.

Figure 7.3 Mean NART errors for participants with ≤11 years education aged 16 – 24 years (n = 27) and participants with >11 years education aged between 61 – 80 years (n = 13).

Chapter 8 – Study 2

Figure 8.1 Mean AUSNART errors according to education for each assessment.

Abstract

The National Adult Reading Test (NART: Nelson, 1982; Nelson & Willison, 1991) is the most commonly used method for estimating premorbid intelligence in neuropsychological practice and research. It has been used extensively in dementia research (Brayne & Beardsall, 1990; Nelson & O'Connell, 1978) and to a lesser degree in traumatic brain injury (TBI) research (Crawford, Parker & Besson, 1988a). Recent findings suggest NART scores may be sensitive to TBI (Riley & Simmonds, 2003), although available studies are limited by small, mixed injury severity samples and ill-defined severity criteria and assessment time post-injury.

Concerns have also been raised regarding the applicability of the NART for Australians as it was developed with a British sample. The Australian National Adult Reading Test (AUSNART: Hennessey & McKenzie, 1994) was developed to address these concerns; however norms were small and limited to Psychology Undergraduate students. There is no published research examining AUSNART in an Australian TBI population.

The aims of the current research were to examine the effects of TBI on NART and AUSNART performance, investigate the influence of demographic variables and compare performance on the two tests. Participants were from a large TBI population study and a longitudinal repeated measures design was implemented.

Study 1 examined the NART performance of 194 participants at 1, 6 and 12 months post-injury. A significant reduction in NART error score was observed by 6 months post-injury, indicating NART scores are sensitive to TBI. A highly significant effect of education was also noted, with higher level of education providing protection from NART 'impairment'. Age, socio-economic status (SES) and TBI severity also significantly influenced NART scores.

Study 2 examined AUSNART performance of 92 participants and was found to be less sensitive to TBI than NART. There was no significant change in AUSNART performance at 6 months, though a significant reduction in error score was noted by 12 months. Age, severity and SES did not significantly affect AUSNART performance but education remained highly influential. AUSNART produced significantly higher estimates of IQ than NART.

Study 3 examined two methods for predicting NART performance following TBI: a Multiple Regression equation using initial NART score and a mean NART change score method based on education and age. Tables are provided for each method to assist clinicians.

Overall, the results indicate caution is necessary when using the NART to estimate premorbid IQ in an Australian TBI population. AUSNART may eventually provide a more valid estimate than NART, although larger normative studies are required before it can be used with confidence.

CHAPTER 1

Overview of the Thesis

The role of neuropsychological assessment in traumatic brain injury has increased significantly since the 1980s when neuropsychological evaluation became regarded as an essential tool of neurological assessment (Kolb &Wishaw, 2009). Individuals with TBI often show no visible sign of cerebral injury on neuroimaging but display significant cognitive deficits in domains such as memory, attention, information processing and executive function. Chapter 2 provides a detailed review of traumatic brain injury research, highlighting prevalence rates, definitions and measurement of TBI as well as a discussion about the outcomes following TBI, according to severity and rehabilitation. Neuropsychologists play an integral role in assessing cognitive deficits following TBI. To judge recovery and for rehabilitation to be effective in achieving the best possible outcome, it is essential to establish the individual's premorbid intellectual functioning to ascertain the degree of impairment. Knowledge of an individual's premorbid IQ shapes expectations of performance on cognitive tests and therefore, determination of the degree of impairment. For example an individual with a premorbid IQ of 130 may be relatively impaired on a memory task but may appear normal if compared to an individual with an IQ of 90. Therefore it is not appropriate to compare an individual's performance to population means as the degree of impairment for the individual can only be ascertained with knowledge of their premorbid intellectual functioning.

Historically, subjective measures were used to estimate premorbid IQ in neuropsychological practice. This relied heavily on the clinician's judgment, comparing prior educational and/or occupational attainment with current cognitive performance. However, these methods were problematic as judgments regarding these variables were misleading, particularly in situations where an individual had not had the opportunity to complete their education fully due to personal circumstances or where psychological disorders, such as depression, anxiety or schizophrenia lead to an inability to be successful academically or to be employed in an occupation they were intellectually capable of fulfilling. In the 1970s the pioneering work of Wilson, Rosenbaum, Brown, Rourke and Whitman was viewed as the first step towards advancing the subjective techniques previously used by clinicians and offered an objective measure of the assessment of premorbid IQ that was independent of current functioning (O'Carroll, 1995). Wilson et al. developed regression equations based on demographic information, such as age, gender, race, education and occupation, which accounted for 54% of the variance in the Wechsler Adult Intelligence Scale (WAIS) Full Scale Intelligence Quotient (FSIQ). This was the catalyst for many studies investigating the validity of the demographic equation in estimating premorbid IQ in countries such as the UK, USA and Australia. These studies are reviewed in Chapter 3. However, this method was also found to be problematic with regression towards the mean, cultural sensitivity restricted IQ range and large standard error of measurement apparent (Vanderploeg, Schinka & Axelrod, 1996).

Subsequently the 'hold' test concept was utilised in the estimation of premorbid IQ. This was based on the premise that certain present abilities, thought to be relatively resistant to neurological impairment (Vanderploeg & Schinka, 1995), could provide an accurate estimate of premorbid IQ. Previously Wechsler (1958) proposed that several WAIS subtests, Vocabulary, Information, Picture Completion and Object Assembly, were minimally affected by the effects of aging and brain impairment and were therefore 'hold' tests while other WAIS subtests, Digit Span, Similarities, Digit Symbol and Block Design, were affected by brain damage and were referred to as 'don't hold' tests. The best-performance method and combined demographic and current measure regression equations became prevalent in research studies utilising the 'hold' test concept. Debate regarding the use of 'hold tests' in the estimation of premorbid IQ developed throughout the 1980s and has continued to date. Some researchers suggest all cognitive tests may be negatively affected by neurological impairment (Langeluddecke & Lucas, 2004) while others state that if considered along with other relevant information this method can be useful and accurate as an estimate of premorbid IQ (Lezak, Howieson & Loring, 2004). These studies are discussed further in Chapter 3.

As objective measures such as demographic equations, 'hold tests' and combined methods were proposed and studied, the development of the NART was occurring. Nelson (Nelson, 1982) developed the NART as an estimate of premorbid IQ in dementia but it soon became the 'gold standard' for the estimation of premorbid IQ for patients with TBI. The acceptance of the NART as the 'gold standard" was based on the premise that the reading of irregular words was a highly practiced and over-learned skill such that once established was maintained despite deteriorations in other areas of intellectual functioning therefore researchers concluded that NART was not sensitive to TBI (Nelson & McKenna, 1975). Chapter 4 provides a review of the development of the NART and Chapter 5 examines its validity with clinical populations such as dementia and TBI.

As the NART became more prevalent as an estimate of premorbid IQ in TBI research and clinical practice, studies emerged challenging the belief that the NART was not sensitive to TBI, sparking concern regarding its generally accepted use. These studies are reviewed in Chapter 6 and highlight several limitations in the field such as small, mixed samples, ill-defined injury severity criteria and assessment time post-injury. There has also been concern regarding the use of the NART in an Australian population as it was developed with British participants. Studies with Australian samples have shown NART accounts for less variance in IQ than for British samples. To address these concerns Hennessey and Mackenzie (1994) developed the Australian version of the NART (AUSNART) with a non-neurological Australian sample and details regarding its development are provided in Chapter 4.

In conclusion, the main focus of the present thesis was to examine the validity of the National Adult Reading Test (NART) as an estimate of premorbid IQ in TBI; to provide a better understanding of the degree of sensitivity to TBI in a large longitudinal repeated measures design. Also, two NART prediction methods are provided to assist clinicians in making more valid estimates of premorbid IQ for patients with TBI. Further to this the thesis examined the AUSNART and validated its characteristics compared to the NART to provide much needed information of this measure as an estimate of premorbid IQ in an Australian TBI population. Chapters 7 – 9 provide details of the three experiments: first, Chapter 7 examines NART performance in an Australian TBI population; secondly, Chapter 8 examines AUSNART performance in an Australian TBI population and a sub-sample who completed both the NART and AUSNART are examined to compare performance on the two tests; and thirdly Chapter 9 examines two methods of predicting the NART following TBI, a regression equation and a mean change score method, according to education and age. Finally Chapter 10 provides a general discussion of the findings of this thesis, summarising the main results and discussing these in relation to previous research findings as well as providing suggestions for future research.

CHAPTER 2

Introduction to Traumatic Brain Injury

Traumatic brain injury (TBI) is the most common cause of brain damage in children and young adults (Lezak et al., 2004). It can cause life-long impairments in physical, cognitive, behavioural and social functioning requiring intensive rehabilitation and support for both the patient and their family (Khan, Baguley & Cameron, 2003).

2.1 Epidemiology

It is difficult to ascertain the true incidence of TBI as estimates are generally based on hospital separations data indicative of incidence but not equivalent to actual incidence rates (Fortune & Wen, 1999). A large proportion of people suffering TBI are not admitted to hospital and many do not present to hospital emergency departments, resulting in an underestimation of the true incidence. Variations in methodology across studies, diagnostic errors and under-reporting of mild head injury significantly affect incidence estimates (Fortune & Wen). However, with this in mind, estimates in 1999 indicated there were approximately 150 people per 100 000 population admitted to hospital in Australia with TBI per year, with severe head injuries accounting for 12 - 14 per 100 000 population, moderate head injuries accounting for 15-20 per 100 000 population and mild head injury thought to account for 70 -85% of all traumatic brain injuries (Fortune & Wen). TBI-associated hospital separations in the 1996-1997 period varied substantially between states with the highest rate for Queensland residents (211 per 100,000) and the lowest for Australian Capital Territory residents (71 per 100,000). A rate of 137 per 100 000 population was observed in Tasmania, resulting in a total of 645 TBI cases in this state in 1999. In Australia the highest age-specific rate was for people aged 15–19 years (284 per 100,000) and 70% of all TBI hospital separations were male. It has been noted that young working class males with limited educational attainment and unstable work history pre injury are representative of the TBI population (Ponsford, Sloan & Snow, 1995). TBI is more common in males than females with a ratio of 3-4:1 and it is thought that the higher proportion of males is related to risk taking behaviour commonly engaged in by young males (Khan et al., 2003).

Methodological differences occur in epidemiological studies resulting in variation in reported statistics. The data used in the current study was from a large populations study at the Neurotrauma Research Register Project (NTR) at the Royal Hobart Hospital (RHH). This data includes all TBI patients who presented to the Department of Emergency Medicine (DEM) regardless of whether they were admitted or not. Conversely, this data does not capture TBI patients who did not present at DEM but choose to visit their GP or who choose not to obtain medical attention. However, this is often the case when attempting to obtain incidence rates. The NTR data includes estimates gathered at one public hospital in Hobart and may therefore mean that the data is more susceptible to local variations in demographic factors such as socioeconomic status (Fortune & Wen, 1999). Notably, the RHH is the only public hospital in Southern Tasmania and services the majority of the population for emergency services.

2.2 Traumatic brain injury defined

Acquired Brain Injury (ABI) is an umbrella term used to refer to brain damage caused by traumatic and non-traumatic events. It is often referred to as the 'hidden disability' because the long term problems in areas of thinking and behaviour are not so easily identified as many other physical disabilities (Brain Injury Association of Queensland [BIAQ], 2002). Definitions and terminology varies significantly across countries and studies (Fortune & Wen, 1999). Non-traumatic ABI refers to stroke, anoxic brain injury, alcohol-related brain injury and brain damage arising before birth, at birth or in childhood. TBI is the most prominent sub-group of ABI and can cause transient or permanent neurological dysfunction (Khan et al., 2003). TBI is defined as:

> "...an insult to the brain caused by an external force that may produce diminished or altered states of consciousness, which results in impaired cognitive states or physical functioning." (Savage, 1995).

The majority of TBIs are termed closed head injuries (CHI) meaning that the skull remains intact and the brain is not exposed. Participants with CHI are examined in the present thesis.

2.3 Types of brain trauma

An external force from a blow to the head with a relatively blunt instrument or from blunt impact with a stationary object can result in brain tissue being torn, stretched, penetrated, bruised or swollen (BIAQ, 2002). The most common cause of head injury in Australia is motor-vehicle (MVA) related, representing approximately two-thirds of moderate and severe TBI. Falls, assaults and sporting accidents are the next most common causes (Khan et al., 2003). Acceleration and deceleration forces, resulting in translation and rotation, can cause laceration of the scalp, shearing of axons, skull fracture and shifting of intracranial contents (Ponsford et al., 1995). Diffuse axonal injury (DAI) is common, particularly following motor vehicle accidents, and is the single most important lesion in traumatic brain injury (Teasdale, 1995). The effect of DAI, even in mild brain injury, is to alter consciousness and the depth and duration of coma, which provides the best guide to severity of the diffuse damage (Teasdale). DAI consists of scattered damage and division of axons throughout the white matter of the brain and is thought to account for much of the disability experienced in the later stages of all TBI (Teasdale). Only approximately 5-10% of DAI cases are visible on Computer Tomography scan (CT), leaving the majority of cases revealing no CT evidence of DAI. However, CT scan is useful in the detection of intracranial haematoma, large contusions, cerebral abscess, ventricular enlargement and atrophy (Ponsford et al., 1995). Magnetic Resonance Imaging (MRI) has been found to be more sensitive to non-haemorrhagic grey and white matter lesions compared to CT, however the greater availability and practicality of the CT make it the most commonly used procedure in the investigation of brain injury (Teasdale).

Focal injuries often occur in the frontal and temporal lobes due to the sharp inner surface of the skull in these areas, often resulting in cerebral contusions at the site of impact, referred to as a "coup" injury and can also occur on the areas opposite the site of impact, referred to as a "contra coup" injury. However contusions may also be found on the medial surfaces of the cerebral hemispheres and along the upper surface of the corpus callosum irrespective of site of impact (Khan et al., 2003; Ponsford et al., 1995). Secondary brain injuries, such as intracranial haematoma, brain swelling, infection, raised intracranial pressure and ischemic damage can occur following complications from the initial injury (Ponsford et al., 1995). Secondary brain injury often occurs as a result of systemic complications and is potentially treatable. Advancing medical technologies have led to many accident victims, who previously would have died, surviving with extremely severe injuries (Lezak et al., 2004). Unfortunately, many remain severely disabled and will, along with their families, confront their disabilities for decades (Ponsford et al.).

2.4 Measuring severity of injury

A major issue in TBI research is the classification of severity of injury. The literature is inconsistent regarding the use of definitions and criteria for severity of injury, which is classified as mild, moderate or severe (Carroll, Cassidy, Holm, Kraus & Coronado, 2004). Loss of consciousness (LOC) and posttraumatic amnesia (PTA) are commonly recognised as measures of severity and outcome following TBI (Ponsford et al., 1995). The Glascow Coma Scale (GCS; Teasdale & Jennett, 1974) is a commonly used measure of depth and duration of impaired consciousness and coma (Ponsford et al., 1995). It was developed to assess duration of coma and to more precisely assess different levels of responsiveness. The scale allows precise recording of different levels of coma in terms of how long different levels of responsiveness have persisted. Three aspects of behaviour are measured independently: motor responsiveness, verbal performance and eye opening (Teasdale & Jennett). These measures are recorded frequently and recordings within the first 24 hours are used to grade severity of injury and predict outcome (Ponsford et al.). A person with a GCS score of 3 - 8 is classified as having a severe head injury, a score of 9 - 12 indicates a moderate head injury and a score of 13 - 15 a mild head injury (Ponsford et al.). A patient is considered to be in a coma if their GCS score is ≤ 8 and they experience coma duration for more than 6 hours after admission (Lezak et al., 2004). The GCS is used in the prognosis for survival and not for functional outcomes (Khan et al., 2003). A study by Williams, Gomes, Owen, Drudge and Kessler (1984) found that estimates of premorbid IQ (based on demographic variables) and coma grade, assessed by GCS, were the most important variables in predicting early cognitive and global outcome. Coma duration has been found to be a poor predictor of outcome for many patients who have periods of coma up to 20 - 30 minutes, but a good predictor of outcome in more severe injuries (Lezak et al., 2004). Teasdale (1995) noted that the definition of mild or minor head injury for a patient with a GCS of 13 - 15 is unsatisfactory as patients scoring 15 make up the majority of patients classified in this group and have a much lower risk of complications at the acute stage and less persistent sequelae than patients scoring 13 or 14. Therefore, grouping all patients with a GCS of 13 - 15 gives the impression of seriousness to those who score 15 and underestimates the severity of those scoring 13 or 14 (Teasdale).

Prior to the 1990s the presence of LOC had generally been essential for a diagnosis of mild TBI. However, a shift in diagnostic criteria from LOC to post-traumatic amnesia (PTA) emerged across disciplines together with a consensus that concussion could be diagnosed even in the absence of LOC (Ruff, 2005). PTA is a period of generalized cognitive disturbance experienced following TBI. A person may be partially or fully conscious although they remain confused and disoriented and are unable to store and retrieve new information (Ponsford et al., 1995). PTA is one of the best predictors of outcome following closed head injury (Ahmed, Bierley, Sheikh & Date, 2000). The duration of PTA includes time in coma and dates from the time of injury until the time the patient is able to store and retrieve new information. PTA duration is often ascertained retrospectively through subjective reporting or through objective measures if a brain injured patient is admitted to hospital. It can be affected by intoxication, pain and medication and further research is needed to explore the potential differentiation between neurogenic and psychogenic gaps in memory (Ruff). While there is some debate as to the accuracy of subjective reporting it is often the only means of ascertaining duration of PTA for patients who are not admitted to hospital. A study by Gronwall and Wrightson (1980) found that in 75% of cases the individual's retrospective assessment of length of PTA agreed with the prospective assessment of the person's level of orientation and capacity to store ongoing memories using an objective measure.

A commonly used objective measure of PTA is the Westmead Post-traumatic Amnesia Scale (Shores, Marosszeky, Sandman & Batchelor, 1986). This scale was developed to provide daily recordings of a person's orientation and ability to lay down new memories. It can be administered daily as a hospital bedside measure and three consecutive days scoring 12/12 indicate the person is out of PTA. The day of the first 12/12 score in the series is recorded as the final day in PTA. The Westmead PTA scale is a preferred measure for assessing length of PTA (Ponsford et al., 1995). When using PTA as a measure of severity the classification proposed by Jennett and Teasdale (1981) is commonly used. This classification is as follows:

- 5 minutes = very mild
- 5 - 60 minutes = mild
- 1 - 24 hours = moderate
- 1 - 7 days = severe
- 1 - 4 weeks = very severe
- >4 weeks = extremely severe

The World Health Organisation (WHO) (Carrol et al., 2004) conducted a review of TBI research and proposed the following PTA severity criteria:

- < 1 day = Mild
- ≥1 day – 1 week = Moderate
- > 1 week = Severe

The Jennett and Teasdale classification of PTA was used in the present thesis as combining participants who experienced minimal PTA (< 1minute) with those experiencing up to one day PTA as proposed by WHO may dilute effects and not adequately assess group differences.

2.5 Outcome following brain injury

The symptoms experienced post TBI are wide and varied and the heterogeneity of these symptoms and the recovery path pose challenges for the patient and their family and for the medical and rehabilitation team assisting in their care. Individual premorbid abilities and psychosocial situations impact significantly on outcome post-TBI, highlighting the need for a holistic, long term and individualised rehabilitation program (Khan et al., 2003). Therefore a valid estimate of premorbid intellectual functioning is vital to the provision of adequate and appropriate rehabilitative care.

Initial consequences and long term outcome post-injury are related to severity of injury and premorbid factors, such as premorbid IQ and psychosocial functioning. A mild traumatic brain injury can involve no loss of consciousness and is often not diagnosed as a brain injury. However, even mild damage to the brain can result in the person responding more slowly and complaining of physical and mental difficulties (BIAQ, 2002). Mild TBI (mTBI) can be classified as loss of consciousness for less than 20 minutes or post-traumatic amnesia for less than one hour (Ponsford et al. 1995). Some common symptoms following mTBI, often referred to as post-concussional symptoms, include fatigue, headache, dizziness, hearing loss, ringing in the ears, concentration and memory problems, sleep difficulties, sensitivity to bright lights, blurred or double vision, anxiety and depression (BIAQ, 2002; Ponsford et al.). The recovery time from these symptoms varies however generally it is expected that the symptoms will subside within a period of days or weeks (Ponsford et al.). The most commonly reported cognitive deficits are memory and attention deficits and it has been argued that deficits initially presenting as memory deficits reflect underlying impairments in attention (Chan, 2000). Reduced speed of information processing leading to a reduced information processing capacity, difficulties in focusing on more than one thing at a time and coping with complexity are the most common attentional difficulties noted in TBI research (Ponsford et al.). Cognitive deficits are thought to resolve within approximately three months for mTBI (Ponsford et al.). Ponsford et al. (2002) developed a booklet, providing psycho-educational information on TBI including coping strategies, and studied the impact of providing this information on outcome following mild TBI. They provided the information one week post-injury and found that it reduced anxiety and lowered the incidence of ongoing problems.

There is increasing evidence suggesting that mild head injury can cause considerable neural damage throughout the brain (Heitger et al., 2004). The study by Heitger et al. found that closed head injury impairs multiple motor systems, particularly motor function originating in the frontal and dorso-parietal brain regions. The authors suggested that abnormalities of saccades and upper-limb visuomotor function following closed head injury may be sensitive markers of cerebral dysfunction.

Moderate and severe TBI involves duration of coma for longer than 20 minutes and PTA for more than one hour and the cognitive and behavioural changes are far more extensive and persistent than in mild brain injury (Ponsford et al. 1995). High incidence of diffuse axonal injury and damage to the frontal and temporal lobes results in a vast array of cognitive and behavioural problems. These include: attentional deficits and fatigue, learning and memory problems, impaired planning and problem solving, concrete thinking, lack of initiative, inflexibility, dissociation between thought and action, impulsivity, irritability and temper outbursts, communication problems, socially inappropriate behaviour, self-centredness, changes in affect and a lack of insight (Ponsford et al.). Many of these cognitive and behavioural problems stem from deficits associated with the frontal lobe and are referred to as executive functions: the capacities for self-determination, self-direction and self-control and regulation dependent on intact awareness of one's self and surroundings (Lezak et al., 2004). Without the capacity of self-awareness many patients are unable to accept the need for rehabilitation, which negatively impacts on social interactions, ability to return to work and driving, and to resume previous activities and relationships (Lezak). Return to work has been used as in index of successful rehabilitation in TBI and the type and duration of rehabilitation has been associated with return to work (McCrimmon & Oddy, 2006).

Olver, Ponsford and Curran (1996) examined the long-term outcome of TBI patients following discharge from a comprehensive rehabilitation programme. They noted that while early intervention achieves significant gains many patients were still making significant functional gains in domestic and community activities between 2 - 5 years post-injury. They suggested that with increasing self-awareness many individuals became more receptive over time to interventions assisting them in returning to work, driving and to community activities, indicating a need for intermittent lifelong intervention for TBI patients (Olver et al.).

2.6 Summary

The assessment of premorbid intellectual functioning of patients with traumatic brain injury (TBI) is essential in establishing degree of impairment and for developing appropriate rehabilitation programs. Inadequate estimates of cognitive decline could potentially alter the course of recovery communicated to the patient and their family and could impact negatively on the number and quality of treatment recommendations (Powell, Brossart & Reynolds, 2003). It is vital therefore that current research examines the validity of estimates of premorbid IQ commonly used as these assist the clinician to establish the extent of impairment for an individual, formulate a maximally effective treatment plan, identify potential issues in readjustment and assist in dealing with legal concerns (Perez, Schlottmann, Holloway & Ozolins, 1996).

With the high incident rates of TBI and a growing body of evidence regarding cognitive deficits, not only following severe injury but also for mild and moderate injuries, it is vital that studies are undertaken to examine the validity of estimates of premorbid IQ specifically for this population. Chapter 3 will review estimates of premorbid IQ commonly used in neuropsychological practice and critique studies conducted to date with normal and clinical populations.

CHAPTER 3

Estimation of Premorbid Ability

The estimation of premorbid ability is a critical component of neuropsychological evaluation (Griffin, Mindt, Rankin, Ritchie, & Scott, 2002). Clinicians are required to make some determination of the extent and pervasiveness of neuropsychological deterioration relative to the individual's pre-injury cognitive functioning (Klesges & Troster, 1987). The diagnosis and assessment of degree of impairment following brain injury or insult requires an estimate of premorbid functioning; for example, a decline in general level of cognitive functioning is required for the diagnosis of dementia (Franzen, Burgess, & Smith-Smeemiller, 1997). Estimates of premorbid intellectual functioning provide invaluable information for the development of rehabilitation programs and in identifying potential difficulties in adjustment to the acquired disability (Perez et al., 1996). In addition, legal issues such as compensation claims following a brain injury require evidence of decline to be clear and convincing and therefore require valid and reliable objective measures of premorbid functioning (Franzen et al.).

Comparing an individual's current performance with group means does not accurately ascertain an individual's prior cognitive status (Klesges & Troster, 1987). This comparison is only appropriate when the score is uniformly present in all individuals and when demographic variables such as age, race, gender or education are not related to performance (Franzen et al., 1997). In most neuropsychological skill areas performance is indeed related to demographic variables therefore suggesting that comparing performance of an individual who has a brain injury with population average scores would lead to misleading inferences (Franzen et al. 1997).

The determination of change in functioning often involves clinical judgment, but using objective measures can greatly assist in reducing the amount of error that can occur with clinical judgment alone (Franzen et al., 1997). Furthermore, Crawford, Millar and Milne (2001b) stated that objective measures can provide unbiased and useful estimates of premorbid IQ to add to the qualitative information available to the clinician. Record of past performance on standardised testing or grade-point average would provide the best and most reliable estimate of premorbid functioning. Unfortunately this information is rarely available and it is therefore necessary for the clinician to rely on current performance measures or demographic equations to estimate the degree of cognitive impairment (Vanderploeg et al., 1996). Numerous methods have been used in the estimation of premorbid ability. These include the use of demographic regression equations, aspects of present abilities measures such as the Wechsler Adult Intelligence Scale (WAIS) Vocabulary subtest (Wechsler, 1981) utilising the hold/don't hold strategy, Lezak's best performance method, combined demographic-current abilities equations, the Oklahoma Premorbid Intelligence Estimate (OPIE) and reading skill (Franzen et al., 1997).

3.1 Demographic-predicted intelligence estimate

Demographic-predicted intelligence estimates are based on known relationships between IQ and demographic characteristics, such as type of occupation or years of education (Crawford et al., 1989d). They have the advantage of being totally objective, do not require judgment by the clinician and are independent of current functioning (Powell et al., 2003). A substantial discrepancy in favour of predicted IQ when comparing predicted IQ with the actual obtained IQ strongly suggests the presence of intellectual impairment (Crawford et al., 1989d).

Several studies have examined demographic variables, such as age, gender, occupation, race and education, as estimates of premorbid IQ and have developed regression equations to aid clinicians. Earlier studies examined the amount of WAIS Full Scale Intellectual Quotient (FSIQ) variance accounted for by demographic variables in a normal population. Wilson et al. (1978) found that demographic variables alone accounted for 54%, 53% and 42% of the variance in WAIS FSIQ, VIQ and PIQ respectively in a US population and Crawford et al. (1989d) found that demographic variables accounted for 50, 50 and 30% respectively for WAIS FSIQ, VIQ and PIQ in a UK population. However, Barona, Reynolds and Chastain (1984), using the US WAIS-R standardization sample, found that demographic variables alone only accounted for 36%, 38% and 24% of the variance of WAIS-R FSIQ, VIQ and PIQ respectively.

Questions regarding the predictive power of this procedure and inconsistent findings highlight some limitations of the use of demographic equations as predictors of premorbid IQ. This method is thought to be culturally sensitive (Green et al., 2008) and regression towards the population mean, restricted IQ range and large standard error of measurement limit the utility of the demographic based regression equations for estimating premorbid IQ (Langeluddecke & Lucas, 2004; Vanderploeg et al., 1996).

3.2 'Hold' / 'don't hold' methods

Estimation of premorbid IQ based on 'hold' tests or on other independent current cognitive measures primarily relies on present abilities which are thought to be relatively resistant to neurological impairment (Vanderploeg & Schinka, 1995). One of the most commonly used 'hold' tests for estimating premorbid IQ is the WAIS Vocabulary sub-test (Lezak et al., 2004). Wechsler (1958) proposed that several WAIS subtests, Vocabulary, Information, Picture Completion and Object Assembly, were minimally affected by the effects of aging and brain impairment, referring to these as 'hold' tests. Several other WAIS subtests, Digit Span, Similarities, Digit Symbol and Block Design, were affected by brain damage and Wechsler labelled these as 'don't hold' tests. However, other research suggests that some 'hold' tests may be negatively affected by neurological impairment (Langeluddecke & Lucas, 2004). A study by Russell (1972) found that all WAIS subtests were negatively affected by brain injury.

A study by Crawford et al. (1988a) found that the National Adult Reading Test (NART; Nelson, 1982; Nelson & Willison, 1991) provided a significantly higher IQ estimates than the vocabulary WAIS subtest for a group of clinical patients, suggesting that the NART was more resistant to cerebral dysfunction than Vocabulary. Notably, there was no significant difference between NART and Vocabulary estimated IQ in the multi-infarct dementia (MID) *(n* = 8) group and the TBI group (*n* = 18), suggesting that both NART and Vocabulary are not impaired in TBI or MID. However, the limited information regarding severity of injury or stage of dementia, length of time post-injury and the small sample size in this study limits the applicability of the findings. Vanderploeg and Schinka (1995) stated that any WAIS-R subtest may be impaired by brain injury and concluded that measures of current ability are problematic in that they may not 'hold' in all types of brain injury or psychopathology.

3.3 Best performance method

Lezak et al. (2004) proposed the 'best-performance method' for the estimation of premorbid intellectual functioning. This method identifies the highest test score in the overall assessment and uses this as the standard to compare all other aspects of the patient's current performance. This method is based on several assumptions put forward by Lezak et al:

> Given reasonably normal conditions of physical and mental development, there is one performance level that best represents each person's cognitive abilities and skills generally..... marked discrepancies between the levels at which a person performs different cognitive functions or skills probably gives evidence of disease, developmental abnormalities, cultural deprivation, emotional disturbance, or some other condition that has interfered with the full expression of the person's cognitive potential..... for cognitively impaired persons, the least depressed abilities may be the best remaining behavioural representation of the original cognitive potential and; within the limits of chance variations, the ability to perform a task is at least as high as a person's highest level of performance of that task (*p*. 97 - 98).

However, Lezak et al. also recommended integrating information from many sources and warned against using a single high test score unless history or observations provide supporting evidence. Mortensen, Gade and Reinisch (1991) noted that while there has been wide acceptance of the 'best-performance method' it has some serious limitations, leading to systematic errors in the expected performance on neuropsychological tests. Mortensen et al. conducted three studies to investigate the validity of the 'best performance method' and concluded that the highest test score should never be used to estimate the premorbid level of general ability as it overestimates premorbid IQ and, furthermore, the lowest test score should never be used to estimate the current level of ability. The first study by Mortensen et al. examined 160 young Danes, free of any neurological disease. Due to the lack of Danish norms the authors formed their own norms based on the results of this sample. They calculated FSIQ (M =100; SD = 15) and to assess the 'best performance method' estimated FSIQ and VIQ by identifying the highest scaled score for each participant. The results revealed that the 'best performance method' significantly overestimated the participants' true intelligence level, with mean estimated IQs of 117.06 and 112.13 for FSIQ and VIQ respectively. A limitation of the first study was the very small age range of participants, yielding a mean age of 23.6 years (SD = 1.4).

Their second study consisted of 120 participants not affected by neurological disease who were administered a large neuropsychological battery. The 'best performance method' was again used to predict performance in the neuropsychological tests from calculated IQs. The aim of this study was to assess how much the mean difference between observed and predicted test performance deviates from zero when the best performance method is used. The results indicated that the best performance method overestimated performance on the standard battery and all subgroups of tests. A third study examined 64 neurological patients with diffuse cerebral atrophy and also confirmed an overestimation of actual intellectual impairment in this group. However, the small age range (*M* = 45, *SD* = 12.92) and low educational experience (*M* = 8.89 years of schooling) of the sample was a limitation in this study.

Lezak et al. (2004) commented on the limitations of the Mortenson et al. (1991) studies noting that the highest WAIS score is not an acceptable predictor of WAIS IQ in cognitively intact participants since the IQ is a mean of all scores both high and low. Using the highest score only will therefore overestimate IQ in intact populations. The authors did not examine other relevant history and information and relied on the test scores alone which deviated from Lezak's recommendations for using this method of estimating premorbid IQ.

It is apparent from the reviewed studies that IQ estimates based on current performance WAIS subtests are susceptible to the effects of neurological impairment (Langeluddecke & Lucas, 2004). The assumption that premorbid performance is comparable to the best performance on any one current measure results in overestimation of IQ and alternate estimates should be considered (Vanderploeg et al. 1996).

3.4 Combined demographic – current measures

The combined demographic-current measure equations also make use of the 'hold' test concept, utilising current scores on tests believed to be insensitive to neurological impairment. These scores are combined with demographic variables to produce equations to estimate premorbid intellectual functioning. While the demographic regression equation method has the advantage of being completely independent of current cognitive functioning, accounting for approximately 50% of the variance in FSIQ, it still leaves 50% of the variance unexplained

(O'Carroll, 1995). Therefore it has been suggested that the addition of current test scores can significantly increase the amount of variance accounted for in FSIQ, VIQ and PIQ (Krull, Scott & Sherer, 1995).

An example of this approach is the Oklahoma Premorbid Intellectual Estimate (OPIE) developed by Krull et al. (1995). This formula was based on combined demographic variables and the Vocabulary and Picture Completion WAIS-R subtests. Krull et al. indicated that these subtests were chosen due to their demonstrated reliability and resistance to neurological insult. Krull et al. divided the WAIS-R standardisation sample into two groups of 940 participants and a premorbid IQ formula was generated using the first group of 940 participants and then cross validated with the second group. The resulting regression equations accounted for 76%, 63% and 75% of the variance of VIQ, PIQ and FSIQ respectively. Demographic variables of age, education, race and occupation were included in all prediction formulas, however, age was not found to contribute significantly to the VIQ equation. The prediction formula for FSIQ included both Vocabulary score and Picture Completion, VIQ included Vocabulary only and PIQ included Picture Completion only. The authors noted that there was a slightly restricted range although it was considerably less than that found using demographic measures alone (Barona et al., 1984). The authors also proposed using either vocabulary or picture completion alone in cases of brain injury where a larger decrement in either verbal or performance processes may occur.

Numerous studies have been conducted examining the OPIE with normal populations (Schoenberg, Scott, Duff & Adams, 2002; Langeluddecke & Lucas, 2004). However several limitations of this method have been observed: first the inclusion of WAIS-III subtests to estimate FSIQ presents potential psychometric issues such as regression to the mean which can lead to over or under estimation of WAIS FSIQ; secondly, high correlations among predictors and criterion leading to multicollinearity, exaggerating the multiple R and decreasing the estimation errors; and finally as some researchers have found all WAIS subtests to be affected by brain injury to some degree then it is unclear how useful this estimate is in TBI.

Langeluddecke and Lucas (2004) compared demographic estimates of premorbid IQ (DP IQ) with the Oklahoma Premorbid Intelligence Estimate-3 (OPIE-3) in an Australian sample including a control group (n = 50), a moderate traumatic brain injured group (n = 35), a severe/ very severe traumatic brain injured group (n = 74) and an extremely severe traumatic brain injured group (n = 41). The OPIE-3 included both demographic variables and current performance measured on selected WAIS-III subtests and the DP IQ was calculated on age, gender, education and ethnicity. The IQ predictions from all OPIE-3 equations showed a significant 'dose-response' relationship to TBI severity. The authors suggested that the findings provided some support for the use of OPIE-3 (Best) for patients with a severe or extremely severe TBI. However, there was less support for its use with patients with a mild to moderate TBI as it overestimated FSIQ by an average of 6 points.

The DP IQ (age, gender, education and ethnicity) yielded average scores across all groups, in keeping with the general population mean. However, the range was severely restricted suggesting that it is only suitable for persons whose premorbid intelligence was likely to be in the average range. For those in the below average range it will overestimate IQ and it will underestimate pre-injury ability in the superior and very superior range. Therefore the use of demographic variables to predict premorbid IQ is not recommended for use with individuals of below average or superior premorbid ability. Langeluddecke and Lucas (2004) concluded that regression to the mean and the high standard deviation for the difference scores (DP IQ - FSIQ) suggests that the DP estimate of premorbid IQ has limited clinical utility.

In summary, the research reviewed indicates that combined current performance measures and demographic variables can account for a significant amount of FSIQ variance in normal populations and can account for significantly more FSIQ variance than either method alone. However, restricted IQ range, regression to the population mean, resulting in an underestimation of IQ at the higher end of ability and an overestimation at the lower end, and high standard error of measurement have all been noted extensively as limitations of demographic equations and combined demographic and current measures as methods of predicting IQ (Langeluddecke & Lucas, 2004). These issues question the utility of such methods when estimating premorbid IQ in a clinical setting. However, it should be noted that regression to the mean is observed for all methods of estimating premorbid IQ (Lezak, 2004). For clinical populations the susceptibility of current performance measures to neurological impairment is likely to result in an underestimate of premorbid ability.

3.5 Reading ability

Word reading tasks, also thought of as 'hold' tests, have been shown to be effective as estimators of premorbid abilities. The use of word reading tests as estimates of premorbid ability were based on four premises (Franzen et al., 1997; Willshire, Kinsella & Prior, 1991):

1. Reading is highly correlated with intelligence in the general population.
2. Reading ability is more resistant to dementia than is the WAIS Vocabulary subtest.
3. The reading of irregular words is more resistant to cognitive decline than regular words.
4. Word reading taps previous knowledge while minimizing the demands on current cognitive capacity.

The application of grapheme-phoneme conversion rules to many English words enables the average literate adult to read regular words correctly irrespective of previous familiarity with, or understanding of, the meaning of the word. However the reading of these words has been found to be impaired in a group of people with dementia compared to a group without dementia (Nelson & O'Connell, 1978). In comparison the application of the grapheme-phoneme conversion rules to irregular words such as *façade* and *psalm* would result in incorrect pronunciation with correct pronunciation dependent upon the subject's familiarity with the word. Therefore the reading of irregular words provides a more sensitive measure of previous familiarity with words rather than a measure of continuing ability to analyse a complex visual stimulus (Nelson, 1982).

Some research has found that reading tests such as the NART are sensitive to neurological impairment from dementia (Stebbins, Wilson, Gilley & Fox, 1990; Fromm Holland, Nebes &Oakley, 1991; Cockburn, Keene, Hope & Smith, 2000) and TBI (Morris, Wilson, Dunn & Teasdale, 2005; Hamish, Godfrey, Harris & Partridge, 2001; Riley & Simmonds, 2003). This suggests caution when using reading tests such as the NART to estimate premorbid IQ. A brief review of reading tests developed for the estimation of premorbid IQ is presented here and a more detailed examination of the NART will be provided in Chapters 4, 5 and 6.

3.5.1 Schonell Graded Word Reading (SGWRT)

The SGWRT (Schonell, 1966) was one of the first popular reading tests to be used for the estimation of premorbid functioning. Nelson (1982) suggested that to establish reading ability as an indicator of premorbid intelligence it would be necessary to show that reading ability is highly correlated with IQ in the normal population and is maintained at, or near, its premorbid level in patients with dementia. Nelson and McKenna (1975) studied this in a group of 98 normal adults and 45 patients with dementia, administering the SGWRT as a measure of reading ability and the WAIS (FSIQ) as a measure of general intelligence. The results showed that word-reading and general intelligence were highly correlated, $r = 0.75$. They also found that word-reading ability was generally well maintained even in patients with widespread dementing processes. Griffen et al. (2002) found a moderate correlation between reading assessed by the Wide Range Achievement Test (WRAT-3) and IQ, $r = 0.63$.

Nelson (1982) noted several limitations of the SGWRT as a measure of reading ability. The SGWRT contains both regular and irregular words and is therefore not a pure measure of a person's familiarity with the word. As mentioned previously the application of the grapheme-phoneme stress rules allow a person to read the regular words without prior familiarity or knowledge of the pronunciation of the word (Nelson & McKenna, 1975). Nelson (1982) also stated that the SGWRT did not contain sufficient difficult items to enable reliable discriminations between the higher IQ levels, producing ceiling effects with a maximum reading score of 100 giving a predicted IQ of 115. The SGWRT contains many long complex words making it very difficult for a patient with dementia to cope (Nelson, 1982). A limitation of this study was the use of only seven WAIS subtests to obtain FSIQ, departing from the standardised administration of the WAIS. As these limitations precluded extensive use of the SGWRT, Nelson (1982) developed the National Adult Reading Test (NART) to address these issues.

Nelson and McKenna (1975) compared reading ability and Vocabulary as indicators of premorbid levels of functioning. The results revealed that discrepancy scores for patients with dementia were considerably higher when based on reading scores compared to vocabulary scores, indicating that reading scores were a better predictor of premorbid functioning than vocabulary. This early study revealed important findings for the use of reading tests in the estimation of premorbid IQ although, as the authors noted, further research with larger samples incorporating variables such as education, occupation and social factors into a multiple regression equation would benefit clinicians who generally are required to make qualitative allowances for these factors.

3.5.2 Wide Range Achievement Test (WRAT-R; WRAT-3)

The WRAT-R was developed by Jastak and Wilkinson (1984) and a more recent version, the WRAT-3, was developed by Wilkinson (1993), although limited research is available on this measure as an estimate of premorbid ability. Johnstone and Wilhelm (1996) studied the clinical utility of reading tests in estimating premorbid intelligence. Using a within-participants design (n = 39) they compared the longitudinal stability of reading, measured by the WRAT-R, versus intelligence, measured by WAIS-R. Due to variability within the diagnostic groups the authors performed separate analysis for three groups based on changes in IQ regardless of diagnosis: a declined group (IQ decline of >0 points over retest); a stable group (IQ improvement of 0 - 6 points over retest) and an improved group (IQ improvement of >6 points over retest). The results revealed no significant change in reading scores over retest for the declined or stable group; however the improved group showed significant improvement in reading scores. This supports the use of the WRAT-R reading test as an estimate of premorbid intelligence for individuals who demonstrate intellectual stability or decline but not for individuals demonstrating intellectual improvement beyond what would be considered practice effects.

While the findings have important clinical implications, several limitations are apparent. The sample was small and included mixed diagnoses of cognitive dysfunction including traumatic brain injury (n = 16), cerebral vascular accident (n = 8) and progressive dementia (n = 2) and 13 patients who had either a medical, neurologic or psychiatric diagnosis. While this was addressed statistically by grouping according to intellectual stability, improvement or decline, rather than by diagnosis it is difficult to apply the findings to specific diagnostic groups including the TBI patients. No information was provided regarding the severity of injury and the mean time post-injury for retest was 28.3 months (SD = 15.2). A range of this magnitude for time post-injury would have an impact on recovery and therefore changes in estimated IQ score. The TBI patients were observed in all three groups, declined, stable and improved. Information regarding severity and time post-injury for retest for each group may have aided interpretation of these findings.

3.5.3 Cambridge Contextual Reading Test (CCRT)

The CCRT (Beardsall & Huppert, 1997) was developed to improve estimation of premorbid verbal IQ in older persons with dementia by placing the NART words in sentences so they were presented in a meaningful context rather than in isolation. Several studies have found the CCRT to be a higher predictor of premorbid intelligence than NART for patients with mild/moderate dementia (Beardsall & Huppert, 1997; Beardsall, 1998). For example, Beardsall (1998) examined NART and CCRT performance for 73 healthy British older persons aged over 70 years. She found that the group as a whole read correctly a mean of 0.9 words less on NART than CCRT. However, when the group was divided into three groups based on number of words read there was a significant benefit of context according to reading ability, with the lowest word reading group showing a mean improvement of three words. The regression equation predicting WAIS-R VIQ from CCRT accounted for 61% of the variance; conversely, Crawford (1992) found 72% of VIQ variance was accounted for by NART. Beardsall suggested this may be due to differences in the sample as Crawford's study had a wider age range and the Beardsall study only examined participants aged over 70 years. Beardsall (1998) found that CCRT was more useful than the NART for predicting verbal intelligence in older persons with low verbal ability. However, low verbal ability may have been due to lower premorbid verbal ability rather than to onset of dementia. In comparison the majority of those with higher reading ability did not benefit from context. There were small numbers in these groups, therefore limiting the generalisability of the findings.

A study by Watt and O'Carroll (1999) supported the use of the CCRT with patients with brain injury and concluded that it was resistant to TBI. However, this study only examined 25 brain-injured patients and used only one measure of severity. Several other studies (Conway & O'Carroll, 1997; Morris et al., 2005) found that by placing the NART words in context (CCRT), performance for more cognitively impaired Alzheimer's patients improved. Morris et al. concluded that while CCRT appeared to be more resistant than NART to brain injury severity, it also appeared to be susceptible to compromise by brain injury. The authors suggested that comparison of estimates of premorbid ability based on the NART, CCRT or WTAR with estimates based upon demographic variables, and consideration of injury severity, may assist in establishing when these measures are likely to have underestimated premorbid IQ. Further research examining the robustness of CCRT to predict premorbid IQ in patients with dementia and other clinical groups, such as TBI, using larger samples is required before the CCRT can be accepted as a useful estimate of premorbid IQ, over and above the NART.

3.5.4 Wechsler Test of Adult Reading (WTAR; Wechsler, 2001)

The WTAR is a 50-item irregular word reading test which has been co-normed, in the UK and US, with the third editions of the Wechsler Adult Intelligence and Memory Scales (WAIS-III; WMS-III; Wechsler, 1997). The WTAR is thought to have the added advantage of offering three methods of estimates of IQ: reading performance; demographic information; and a combined reading performance/demographic equation, with the combined method producing the lowest prediction error and recommended as the most accurate estimate of IQ (Mathias, Bowden & Barrett-Woodbridge, 2007). Mathias et al. compared NART and WTAR IQ estimates with concurrent WAIS-III performance in a sample of 93 healthy Australian adults. Both WTAR and NART estimates of IQ were highly correlated with concurrent WAIS-III VIQ and FSIQ ($p <.001$), although a combined WTAR and demographic variable equation (U.K.

norms) was found to be a significantly better predictor of WAIS-III VIQ (*r* =.81) and FSIQ (*r* = .78) accounting for 65% and 60% of the variance in VIQ and FSIQ respectively. In comparison the NART accounted for 51% and 44% of the variance in VIQ and FSIQ, the WTAR alone accounted for 53% and 49% of VIQ and FSIQ variance and demographic variables accounted for 42% and 37% of the variance in VIQ and FSIQ respectively. Examination of the difference scores between estimated VIQ and FSIQ with concurrent VIQ and FSIQ showed that all estimated IQs significantly underestimated concurrent WAIS-III IQ. However the NART estimate of VIQ was found to be the best estimate revealing the smallest amount of difference with concurrent WAIS-III VIQ (mean difference estimated - concurrent = -2.3, *SD* =9.6). The results also showed that the WTAR and the NART both overestimated IQs below 100 and underestimated IQs over 100 and it was noted that in some cases IQ was overestimated by up to 30 points and underestimated by up to 36 points.

Comparisons of WTAR reading, demographic estimates and combined demographic-current abilities estimates of IQ (VIQ & FSIQ), calculated using both U.S. and U.K. norms, with concurrent WAIS-III IQ were examined. The WTAR estimates of IQ calculated with both the U.S. and U.K. norms were not significantly different from one another and both IQ estimates were highly correlated (VIQ: $r = .95$; FSIQ: $r = .96$) and significant ($p<.001$). The authors concluded the accuracy of the NART and WTAR as an estimate of premorbid IQ is affected by the person's level of ability with underestimation of IQ in individuals whose IQ is above the population mean of 100 and overestimation of IQ for individuals whose IQ is below 100. Therefore these estimates should be assessed with other sources of information about premorbid ability which will provide information regarding the applicability of this method of estimating IQ for a particular individual. The sample studied had a limited age range and educational level consisting of a predominantly young sample with a mean IQ falling in the high average range suggesting that studies with a broader age range and educational background are required.

A study by Mathias, Bowden, Bigler and Rosenfeld (2007) examined the validity of the WTAR as an estimate of premorbid IQ for patients with TBI. They examined mild (n = 21), moderate (n = 31) and severe (n = 26) TBI groups and a control group (n = 21) and administered the WTAR 3 - 6 months post TBI and again approximately 6 months later. The results showed a significant group difference, with the severe TBI group showing significantly lower WTAR scores (four points) than all other groups, suggesting that word reading is affected by severe TBI. The other groups did not differ from one another. There was a significant improvement in WTAR scores over time but this was minimal, with an improvement of approximately one point only. The absence of a group by time interaction indicated there were no group differences in the extent to which WTAR score improved over time (Mathias et al., 2007). As the authors suggest, there were small changes over time for all groups which may reflect practice effects rather than recovery from TBI. The severe TBI group appeared to have suppressed WTAR scores both at initial and subsequent testing with a gain of only one point over time. However, there were only small samples for each group resulting in a loss of statistical power. These findings were, however, supported in a study by Green et al. (2008), also observing stable performance on the WRAT at 2 and 5 months post-injury for 24 patients with severe TBI. Riley and Simmonds (2003) observed significant improvements in NART error scores up to four years post severe TBI, suggesting the severe group in the

Mathias et al. (2007) study and all participants in the Green et al. study may require more time for recovery and subsequent improvement in WTAR scores to be observed.

3.6 Lexical decision making task

The Spot-the-Word Test (STWT) is a lexical decision task designed to provide a robust estimate of verbal intelligence. It was developed by Baddeley, Emslie and Nimmo-Smith (1993) to provide a supplement to the NART with the aim of addressing NART limitations. The STWT is a lexical decision task requiring the patient to choose the real word when presented with two words, a real word and an invented non-word. The task can be presented in the auditory or visual modality and a number of parallel routes are thought to be involved in performing the task. The routes involved can be based upon the meaning of the word, its orthographic appearance, its sound and a general feeling of familiarity with the word, suggesting that it would be more resistant to brain damage than a task based on a single feature (Baddeley et al.).

These authors proposed that the STWT addressed several limitations of reading tests such as the NART. These include:

1. The inability of the NART to be successfully administered to patients with dyslexia, visual or articulatory problems.
2. Underestimation of the NART when used with patients who are self-educated and have acquired knowledge through reading, but may not be familiar with the pronunciation of low-frequency words despite knowledge of their meaning.

3. The NART cannot be adapted to other languages due to the irregular orthography of the words.
4. The degree of difficulty of NART words may be discouraging to some patients.

The authors provided reliability and validity information gathered over three studies. STWT was found to be highly correlated with NART (*r* = .83) and revealed good validity and reliability. The authors concluded that the STWT was a useful additional method of estimating premorbid intelligence although there are limited studies on its utility as a measure of premorbid IQ and further studies are required to explore its robustness.

3.7 Summary

Historically, the estimation of premorbid IQ in neuropsychological practice consisted of subjective methods relying heavily on clinical judgment and on occasion records of past performance, although this was rarely available. The 1970s saw the introduction of objective measures such as demographic regression equations to estimate premorbid IQ followed by the 'hold' test concept based on the assumption that some WAIS subtests such as Vocabulary were resistant to brain injury. Regression equations combining demographic variables and 'hold' test scores were developed to increase the amount of variance accounted for in WAIS FSIQ. Despite these objective methods providing assistance to clinicians when estimating premorbid IQ they were limited by regression to the mean, cultural sensitivity, multicollinearity, restricted IQ range, large standard error of measurement and they did not take into account the susceptibility of hold tests to brain damage.

While the use of reading tests for the estimation of premorbid IQ is also based on the hold test concept, many studies provided evidence to support the assumption that reading ability is a good indicator of premorbid intelligence even in the face of cognitive decline (Nelson & McKenna, 1975; Griffen et al., 2002; O'Carroll & Gilleard, 1986; Sharpe & O'Carroll, 1991). However, given the findings that 'hold' tests previously thought to be resistant to cognitive decline or brain injury may be sensitive to TBI it will be important for research to explore the validity of reading tests as estimates of premorbid IQ in TBI.

Chapter 4 will discuss the development of the NART and its clinical utility as an estimate of premorbid intellectual functioning, examining its ability to estimate premorbid IQ in normal samples. Studies examining the validity and reliability of the NART for use with individuals with dementia and TBI will be critiqued in subsequent chapters.

CHAPTER 4

Development of the National Adult Reading Test (NART)

Nelson (Nelson, 1982; Nelson & Willison, 1991) developed the NART to assist clinicians in estimating premorbid IQ in patients with dementia. It is one of the most commonly used measures to estimate premorbid intelligence (Crawford, Deary, Starr & Whalley, 2001a). Nelson (1982) proposed that the NART was a sensitive measure of previous familiarity with words rather than a measure of continuing ability to analyse a complex visual stimulus and therefore a useful measure for estimating premorbid IQ. Reading is thought to be a highly practiced and over-learned skill and once established it can be maintained despite deteriorations in other areas of intellectual functioning (Nelson & McKenna, 1975).

Prior to the development of the NART the WAIS vocabulary subtest, referred to as a 'hold' test, had been used extensively as an estimate of premorbid IQ, it had been found that patients with dementia appeared to read surprisingly well, indicating that previous familiarity with a word was not negatively affected by cognitive decline (Nelson, 1982). However, it was necessary to show that reading ability was highly correlated with general IQ level in the normal population and that reading ability was maintained at or near its premorbid level in patients with dementia before word-reading ability could be widely accepted as a measure of premorbid IQ. Nelson and McKenna (1975) examined the relationship between WAIS FSIQ and reading ability, assessed by the Schonnell Graded Word Reading Test (SGWRT), and observed a high correlation ($r = 0.75$). This correlation was only marginally lower than the correlation for reading score and Verbal IQ ($r = 0.78$). The authors suggested this indicated a regression equation based on FSIQ did not present any significant loss of accuracy in prediction compared to the use of an equation based on Verbal IQ only. Other studies have also observed high correlations, ranging from 0.74 to 0.81, between IQ and NART, (Blair & Spreen, 1989; Crawford et al., 1989a; Nelson & O'Connell, 1978; and Sharpe & O'Carroll, 1991).

Nelson and McKenna (1975) also demonstrated that word-reading ability was well maintained even in dementing patients. The authors found while the WAIS Vocabulary subtest produced the highest mean age-corrected WAIS subtest score for the patients with dementia this was still significantly lower than the mean WAIS Vocabulary score for the control group. In comparison the mean reading score for the patients with dementia and the control group were similar, providing further confirmation for the use of word-reading in preference to the WAIS Vocabulary subtest when estimating premorbid IQ. They also observed a smaller overlap between controls and patients with dementia for the discrepancy score when it was based on reading compared to Vocabulary. However this study was limited to a small heterogeneous sample ($N = 45$) of dementia patients compared to a larger control group ($N = 98$). Ceiling effects were noted and the SGWRT revealed several limitations as a word-reading test, particularly in regards to its inclusion of both irregular and regular words and its inability to differentiate between the higher levels of intelligence because of its ceiling level of 100 words yielding an equivalent of 115 for FSIQ.

In response to the limitations of the SGWRT, Nelson developed the New Adult Reading Test (Nelson & O'Connell, 1978) and in 1982 Nelson published the renamed National Adult Reading Test manual. It was designed specifically to provide a means of estimating premorbid intelligence in adults with suspected intellectual deterioration. The NART is a 50 item phonetically irregular word list thought to be relatively unaffected by neurological and psychiatric disorders (Crawford et al., 2001a). The words are presented in order of increasing difficulty and they cannot be pronounced by common rules of pronunciation such as phonetic decoding. The subject reads the words aloud and error scores are used to estimate WAIS FSIQ, VIQ and PIQ using the tables provided in the manual (Nelson & Willison, 1991). Discrepancy scores between predicted IQ and obtained IQ provide information regarding the extent of intellectual deterioration. Tables are also provided detailing the distribution of predicted–obtained IQ discrepancies in the general population (Nelson & Willison). A positive predicted–obtained IQ discrepancy indicates that the individual may have functioned at a higher level pre-injury. The larger the positive discrepancy the more likely intellectual deterioration has occurred (Nelson, 1982). Crawford et al. (1989d) observed that a discrepancy in favour of predicted IQ of more than 15 points occurred in only 1% of an unimpaired UK sample (N = 151) and noted that such a discrepancy strongly suggested intellectual impairment.

The standardization of NART (Nelson, 1982) included 120 British participants who were inpatients with extra-cerebral disorders at the National Hospital for Nervous Diseases. Participants ranged in age from 20 - 70 years and were stratified into five social classes. A short WAIS was administered consisting of seven subtests of the WAIS and the SGWRT. Nelson and Willison (1991) observed a high correlation between the full WAIS and the shortened WAIS (seven subtests) used in this study (r = 0.98) suggesting that the use of the shortened version is adequate for obtaining IQ scores in a normal population. Conversely, Crawford et al. (1989a) suggested that the Full WAIS-R should be used in preference to the short WAIS-R as they found greater predictive accuracy of NART IQ estimates for the full WAIS than the prorated WAIS in a cross-validation sample. The original standardisation sample (Nelson, 1982) yielded an above average mean IQ level and the higher social classes were over-represented. The authors noted there was no significant difference between mean predicted IQ and mean obtained IQ within each social class and therefore concluded that social class did not have a significant independent effect on word reading ability (Nelson & O'Connell, 1978). A limited range of obtained IQ scores (86-128) and ceiling and floor effects for NART predicted IQs were observed. The authors suggested caution when using the NART to estimate IQ for individuals with a premorbid IQ above 125 or below 80. The NART equation accounted for 55%, 60% and 32% of WAIS FSIQ, VIQ and PIQ variance respectively.

Age did not correlate significantly with the number of NART errors, confirming previous findings suggesting age does not influence current word-reading ability in normal participants and patients with dementia (Nelson & McKenna, 1975). Crawford, Stewart, Garthwaite, Parker and Besson, (1988b) found that a curvilinear relationship did not exist between age and NART performance, suggesting a correlational technique was permissible. They found that age did correlate with NART estimated IQ as did education and social class. After partialling out the effects of education and social class the correlation for age was no longer significant. The gender ratio and the effect of gender on NART performance was not reported in the standardisation sample (Nelson, 1982), although Crawford et al. (1988b) found no significant effect of gender on NART performance. Crawford et al.'s study used a larger sample consisting of 201 UK participants ($F = 129$; $M = 72$) free of neurological or psychiatric disorder and a wider age range (17-88) than the standardisation sample. Notably there was an under-representation of social classes 4 – Partly-skilled occupations (12.5%) and 5-Unskilled occupations (3.1%) which may have influenced the effect of social class on NART performance.

Nelson and Willison (1991) produced a restandardisation of the NART against the WAIS-R, as the IQ equivalents of NART scores given in the original standardisation were applicable to WAIS IQ only. Wechsler (1981) found that the WAIS-R gave IQs, which were on average seven and a half points lower than the original WAIS. Therefore it was necessary to develop norms and produce equations for NART predicted WAIS-R IQs. The restandardisation sample were 182 British participants (*M* = 92; *F* = 90) without neurological deficits with an age range of 18 - 70 years. There were participants representing all five social classes (social classes 1 – Professionals, and 5 – Unskilled occupations were under-represented) and the range of WAIS-R IQs predicted from NART were 73 - 131. The obtained WAIS-R IQs ranged from ≤80 to ≥130. Notably, the NART is a relatively poor predictor of PIQ and while PIQ is commonly more affected by neuropsychiatric conditions it may also be that predicted-obtained PIQ discrepancies have lower discriminative ability than FSIQ or VIQ (Crawford et al., 1989a).

Nelson and Willison (1991) examined the effect of social class on predicted – obtained FSIQ discrepancy and found no significant effect of group. Once again there was no significant effect of age on NART score and age was not correlated with IQ. There was a significant difference between males' and females' predicted-obtained FSIQ, however, the authors noted that adjusting scores to allow for gender effects had little effect on the overall accuracy of prediction.

In an attempt to address the issue of ceiling and floor effects Nelson and Willison (1991) first added a set of ten hard words to see if this would extend the range of predictable IQs. This did not substantially increase the upper limits of the test, although the authors indicated that the standardisation sample included more high IQ participants to address the problem of underestimating the IQ of bright participants. In another sub-study (n = 92) they added ten easy words in an attempt to extend the range of predictable IQs downward. The results indicated very few errors were made on these words by participants with more than 10 correct NART words but the addition of these words extended the IQ range down to approximately 55, although further studies examining the ability of the ten easy words to extend the range of predicted IQs are required.

Large studies on neurological samples with specific populations, such as TBI, are required to thoroughly examine the effects of demographic variables such as age, gender and occupation on NART performance.

4.1 Reliability

Research indicates that NART is a reliable test for psychological practice and research. Split-half reliabilities of 0.90 and 0.93 have been reported in previous studies (Crawford et al., 1988b; Nelson, 1982). O'Carroll (1987) studied the inter-rater reliability of the NART and found it had high inter-rater reliability (0.96 – 0.98). Unfortunately this study was very small (*n* = 12) and the sample was within a very narrow range of intelligence (FSIQ range: 102 - 125). Further studies with larger samples and a wider range of NART performance suggest that not only does the NART possess high inter-rater reliability but it can be reliably administered by inexperienced as well as experienced clinicians (Crawford et al., 1989a; Riley & Simmonds, 2003). Schlosser and Ivison (1989) also found high inter-scorer reliability for the NART in a normal group (0.98) and a dementia of the Alzheimer type (DAT) group (0.97). While they did not give details of the number or experience of raters they indicated that the NART was scored immediately for the total group ($N = 81$) and tape recorded to be scored by an independent rater later.

Crawford et al. (1989a) provided a more detailed item analysis of the percentage agreement between raters for individual NART words in preference to examining total scores. They found that 82% of the words had 90% agreement and 64% had a ≥95% agreement. They noted that care should be taken when recording and scoring the words: *aeon, puerperal, aver, sidereal* and *prelate* as they yielded the lowest agreement rates. A study by Alcott, Swann and Graham (1999) also confirmed the finding of low inter-rater reliability for the words listed but noted that training in the pronunciation of these words increased their reliability. Thus with training in the pronunciation and administration of NART words, inter-rater reliability is good.

Test-retest reliability has also proven to be extremely high ($r = 0.98$) for the NART in a study by Crawford et al. (1989a). There was a significant decrease in NART error score at re-testing, however, the mean change was less than one NART error with the authors suggesting practice effects are of little practical significance.

4.2 Validity

The NART has been shown to provide valid premorbid estimates of general intelligence in the normal population (Crawford et al., 1989a; Crawford et al. 1989b; Crawford, Stewart, Cochrane, Parker, & Besson, 1989c; Crawford et al., 2001a; Crawford et al., 1990b; Berry et al., 1994). Crawford et al. (1989c) found that NART loaded highly on '*g*' (0.85), the general factor of intelligence, in a UK sample of 139 (M = 70; F = 69) participants free from neurological, psychiatric or sensory disability.

Another study by Crawford et al. (1989a) demonstrated that NART was a valid estimator of premorbid IQ in a larger sample consisting of the combined standardisation sample (n = 120) and a cross-validation sample (n = 151). The cross-validation sample had a wider IQ range than the standardisation sample (75-140 vs. 86-128) and a larger age range (16-88 vs. 20-70 years) and NART equations accounted for 66%, 72%, and 33% of the variance in WAIS FSIQ, VIQ and PIQ respectively. Both samples were combined (N = 271) yielding equations, which accounted for 57%, 63% and 31% of the variance in FSIQ, VIQ and PIQ respectively. The authors suggested that the slightly lower predictive accuracy for the combined sample may have been due to the use of the prorated WAIS rather than the full WAIS and recommend the full WAIS be administered in further validation studies.

Berry et al. (1994) conducted an important study assessing the validity of the NART retrospectively using the New Adult Reading Test-Revised (NART-R). The NART-R was developed by Blair and Spreen (1989) to provide modifications for acceptable pronunciations for American and Canadian participants. This study was the first to examine the validity of the NART-R to estimate WAIS-R IQ scores obtained 3.5 years earlier and to provide a cross-validation of the NART-R in a sample of older Americans. The sample consisted of 54 participants ($F = 27$; $M = 27$) from the US with a mean age at baseline of 67.8 years (SD = 8.6). NART-R estimated WAIS-R FSIQ, VIQ and PIQ were highly correlated with WAIS-R FSIQ ($r = 0.70$), VIQ ($r = 0.68$) and PIQ ($r = 0.61$) obtained 3.5 years earlier. However the differences between mean obtained and estimated FSIQ (-3.8 points) and VIQ (-5.3 points) were statistically significant. PIQ did not reveal significant differences between estimated and obtained scores. The authors refer to Blair and Spreen (1989) who suggested that a difference of ≥ 15 points in favour of estimated IQ is required for a significant decline in cognitive ability to be noted. If estimated IQ is consistently several IQ points below the obtained IQ as shown in the normal sample, then it could be expected that an underestimate of the actual cognitive decline in a neurologically impaired individual would also be observed.

However when the WAIS-R scores were regressed on NART-R error scores to examine predictive equations specific to the sample studied, the differences between estimated and obtained IQs were reliably smaller than the differences observed using the original Blair and Spreen (1989) regression formula. Berry et al. (1994) noted that the increased accuracy using their formula may have been due to the variance specific to this sample and cross-validation is required before it can be used with confidence. The study consisted of older participants and some research suggests that decline in intelligence is common in this age range (Berry et al., 1994). However Snow et al. (1989) examined test-retest reliability of the WAIS-R FSIQ in normal older Canadians at a one-year interval and found it to be very high (0.90). The lack of WAIS-R data at follow-up in the Berry study was a weakness and a more thorough examination of the presumed stability of intelligence and WAIS-R scores over time would have been permitted if the WAIS-R had been administered at follow-up.

Crawford et al. (2001a) also provided retrospective validity of the NART in a study of 179 individuals who had taken an IQ test (The Moray House Test [MHT]) at the age of 11 years. They administered the NART and the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) to these individuals at age 77 and a MHT equivalent IQ test was re-administered to a sub-set of 97. The MMSE is a screening assessment for cognitive impairment. Crawford et al. stated that if the NART is primarily an index of prior intelligence then it should correlate as highly or even more highly with earlier IQ scores than with current IQ scores. Also, the correlation between earlier IQ scores and NART would be as high, or possibly higher, than the correlation between the two sets of IQ scores. Using a standard formula the authors corrected for the effects of restriction of range resulting in a highly significant correlation between NART at age 77 and the MHT at age 11 of $r = 0.78$ in the full sample, with the NART accounting for 61% of the variance in IQ. The correlation between NART and MHT in 1932 ($r = 0.69$) for the sub-sample was similar to the full sample and was higher, although not significantly, than the correlation between NART and current MHT ($r = 0.63$). Furthermore, the correlation between NART and MHT scores from the 1932 testing ($r =$

0.69) was comparable to the correlation between the MHT scores administered at two different time points, 1932 and 1998 ($r = 0.64$).

Crawford et al. (2001a) also assessed the relationship between NART and MMSE. The authors proposed that if the relationship between NART and MMSE stems from a shared sensitivity to cognitive decline then statistically controlling for childhood intellectual ability should leave the correlation unaffected but if the shared variance was due to the influence of prior intellectual ability then controlling for childhood IQ should weaken the relationship. The results yielded a significant correlation between NART and MMSE ($r = 0.25$), however, when childhood IQ was controlled for, the correlation was no longer significant ($r = 0.02$) and the partial correlation was significantly lower than the raw correlation. The results supported Crawford et al.'s (2001a) proposal that performance on an impairment sensitive test is partly determined by prior IQ and does not reflect individual differences in age associated cognitive decline.

There were several limitations in this study, firstly, as the authors note, the relationship between NART and MMSE may vary in different populations. The sample used in this study was non-clinical and it is possible that in a different population the relationship could vary and be due to a shared sensitivity to impairment, particularly in severely impaired populations as has been suggested in previous studies (e.g., Fromm et al., 1991). The MMSE had limited variance in this sample and scores were within the normal range of 24 - 30 with ceiling effects noted. The authors indicated that the test given in 1932 was not referred to as the Moray House Test No 12 but closely resembled it. However it is difficult to assess the comparability of these tests, and it is notable that for the sub-sample studied the mean MHT-1932 score was 43.20

compared to a mean score of 54.01 for the MHT-1998, suggesting an increase in mean scores and therefore IQ over time, or the possibility that the tests administered in 1932 and 1998 were not directly comparable.

To examine the ability of the WAIS and NART to discriminate between healthy participants and patients with either dementia or cortical atrophy Crawford et al. (1990b) employed a hierarchical discriminant function analysis. In comparing the dementia sample (n = 24) with the healthy sample (n = 151) the addition of NART increased the percentage of cases correctly classified from 85.8% to 96.2% and for the atrophy sample (n = 8) compared with an increase in the healthy sample (n = 151) from 78% to 87.4% for FSIQ. The authors suggested that this was due to the NART's ability to partial out any unwanted variance in the WAIS measure rather than its ability to directly discriminate between impaired and non-impaired participants. The clinical samples were small in this study, particularly the atrophy group, consisting of eight participants only and further studies with larger clinical samples are needed to confirm these findings.

4.3 NART & WAIS-R validation

To date there have been limited validation studies for the use of NART with WAIS-R in normal populations. The Blair and Spreen Revised New Adult Reading Test (1989) developed for North American and Canadian participants (N = 66) has been validated with the WAIS-R. They found that NART-R accounted for 56%, 69% and 16% of the variance for FSIQ, VIQ and PIQ respectively and the addition of demographic variables did not significantly increase the amount of variance explained. The authors noted that the limited variability in education and occupational status in their sample may have reduced the predictive accuracy of demographic variables. However, as the NART-R was standardized and validated with the same sample further validation studies are required with normal populations.

Wiens, Bryan and Crossen (1993) replicated Blair and Spreen's (1989) study, examining the validity of NART-R (Blair & Spreen, 1989) with the WAIS-R. They also provided normative data studying a large sample of country dwelling US individuals (n = 302) and they examined the relationship between NART-R and a commonly used word-reading test, the Wide Range Achievement Test (WRAT-R). The age range was 20-54 years and there were 241 males and 61 females. The sample was relatively young and the correlation between age and NART errors was insignificant (r = .04). There was a significant but very modest correlation between NART errors and education (r = -.14). They found no significant gender differences on NART error scores or WAIS-R IQ scores. There was a significant correlation between NART-R errors and obtained WAIS-R FSIQ (r = - 0.46) and between NART-R errors and VIQ (r = - 0.56). These correlations were considerably lower than those observed by Blair and Spreen (1989) (-0.75 and -0.83 for FSIQ and VIQ respectively). However, the Blair and Spreen (1989) equation fairly accurately predicted FSIQ for participants within the average intellectual range but it overestimated FSIQ at the lower end of the range and underestimated IQ at the higher end. For participants who obtained FSIQ in the 80 - 89 range the NART-R equation overestimated by approximately 20 points, in the 90-100 range an over estimation of 8 points was observed. Conversely, in the 110-119 range the equation under-estimated FSIQ by 8 points and for those with obtained FSIQ of 120 and above it under-estimated by 15 points. The authors question the value of the NART-R to estimate premorbid IQ in individuals who are outside the average intellectual range.

Grober and Sliwinski (1991) also examined the validity of NART to predict WAIS-R IQ. They used the American version of the NART known as the AMNART, developed by Schwarz and Saffran (1987). It was standardized on 109 participants without neurological deficits, aged between 40 - 89 years. The AMNART consists of 50 words in total, with 23 NART words which were unfamiliar to American participants replaced by irregular American words. Grober and Sliwinski (1991) examined 230 participants who did not have a diagnosis of dementia and they divided this group into two to enable development and cross validation of their model. They developed a model to predict premorbid IQ based on AMNART error scores and years of education. They applied the model to a sample of 25 participants with a diagnosis of dementia. An abbreviated version WAIS-R was administered and used to compute a prorated score for current VIQ. They found that estimated VIQ for the patients with dementia did not differ from that estimated for the non-dementia participants. However, the patients with dementia displayed a discrepancy of at least 10 points, with estimated VIQ exceeding obtained VIQ. The authors suggested that this method could be used to identify the presence of intellectual decline. However, they also note that there is a need to apply this model to an unselected sample of elderly people to determine its sensitivity and specificity to distinguish demented form non-demented participants. They also question its applicability to people with advanced dementia as a significant correlation between dementia severity and premorbid IQ was observed, indicating the model may underestimate IQ in persons with moderately severe dementia. Therefore further studies are required to examine how the model is affected by dementia severity.

4.4 NART & WAIS-III validation

Validation studies for the use of NART with WAIS-III (Wechsler, 2002) are not available currently. However, Crawford developed provisional norms for the use of NART with the WAIS-III which were provided to the supervisor of the current author and used in the present thesis (Crawford, 1997). Further research is required to provide validation of the NART with WAIS-III particularly in light of the extensive changes to the structure of the WAIS-III.

4.5 Validation of NART with clinical groups

Early validation studies of the NART with clinical groups suggested the NART was a valid and useful tool for predicting premorbid IQ. Nelson and O'Connell (1978) studied 40 patients from the National Hospital whose records showed evidence of bilateral cortical atrophy and a control group consisting of the 120 participants used in the NART standardisation study. They compared the ability of NART ('irregular words') and the SGWRT ('regular words') to predict premorbid intellectual functioning and found the NART to be a more sensitive measure for patients with dementia. They found that mean NART error scores were not significantly different for the atrophy group ($M = 23.9$, $SD = 11.2$) than the control group ($M = 22.4$, $SD = 10.1$) despite significantly lower mean FSIQ for the atrophy group ($M = 91.6$, $SD = 15.9$) compared with the control group ($M = 109.2$, $SD = 11.3$). The obtained IQs were consistently poorer than the NART predicted IQ for the atrophy group indicating that the reading of irregular words was relatively well preserved despite cognitive decline. However, as noted previously, the study consisted of a small heterogeneous sample of dementia patients.

Crawford, Besson, Parker, Sutherland and Keen (1987) examined the validity of the NART and the Vocabulary subtest of WAIS to predict premorbid IQ in patients with depression (n = 39) and matched controls, given that depression can cause cognitive impairment. NART provided a higher estimate of IQ (M = 107.4, SD = 10.2) than the Vocabulary sub-test (M = 102.3, SD = 13.2) in 79% of patients with depression. There was no significant difference in NART performance between controls and patients with depression; however, there was a significant difference in performance on the Vocabulary sub-test between the controls and depressed patients indicating that NART 'holds' in depression.

Another study by Crawford et al. (1988a) examined the validity of NART as a measure of premorbid IQ in a mixed group of neurological patients (N = 70). They found that while NART and Vocabulary estimated IQ did not differ significantly in the control group, the NART provided a significantly higher IQ estimate than Vocabulary for the clinical sample. This suggests that the NART was more resistant to cerebral dysfunction than Vocabulary. NART performance for patients with alcoholic dementia, dementia Alzheimer type (DAT), multi-infarct dementia (MID), and TBI did not differ from matched controls, again supporting its use in these clinical groups. Korsakoff and Huntington's groups performed significantly lower on the NART than their matched control groups and its use with this patient group is questioned. The TBI group was the only clinical group that did not perform lower on the Vocabulary test than their matched control group, suggesting that Vocabulary may also be a valid measure for estimating premorbid IQ in patients with brain injury. However, the authors did not provide information regarding severity of injury or time post-injury for the closed head injury group and the study was limited by the very small number of participants in each clinical group.

A small Canadian study by Sharpe and O'Carroll (1991) examined NART and WAIS-R Vocabulary performance in elderly patients with dementia ($N = 20$) and in matched controls ($N = 20$). They found that WAIS-R FSIQ and VIQ were significantly higher for the control sample ($M = 96.9$ $SD = 13.2$ & $M = 98.0$, $SD = 15.1$ respectively) than the patient sample ($M = 74.5$, $SD = 9.1$ & $M = 77.7$, $SD = 8.7$ respectively). NART estimated FSIQ ($M = 88.85$, SD = 10.7) was significantly higher than Vocabulary estimated IQ ($M = 84.85$, $SD = 8.38$) in the patient sample and both estimates were significantly higher than the obtained WAIS-R IQ ($M = 74.75$, $SD = 9.14$) for this group. However NART error scores for patients ($M = 35.2$, $SD = 10.7$) were not significantly higher than those obtained by controls ($M = 27.3$, $SD = 10.2$). The authors concluded that NART remains relatively unaffected by the dementing process. Once again, the sample size was small and further validation studies with larger samples are necessary.

Beardsall and Brayne (1990) studied a UK elderly female community sample ($N = 316$) including normal, mild dementia and mild/moderate dementia participants, divided into two age ranges, 70-74 years ($n = 185$) and 75-79 years ($n = 180$). They examined the effect of education based on school leaving age (≤ 14 years and > 14 years). There was a significant increase in predicted IQ for those with more education. They found no effect of age within each education group for predicted IQ confirming previous findings of no relationship between age and NART performance (Nelson, 1982; Nelson & Willison, 1991; Crawford et al. 1988b). Further analyses were performed comparing NART predicted IQs for three groups, normal, mild dementia and mild/moderate dementia within each age group. There was a significantly lower NART predicted IQ for participants with dementia (Mild: $n = 97$; Mild/Moderate: $n = 99$) compared to controls ($n = 104$) but the mild/moderate group did not yield significantly lower NART predicted IQs than the mild dementia group as might be expected.

The authors concluded that the NART was an acceptable measure of premorbid intelligence for most people in the community. They suggested that the significant difference in predicted IQ's between controls and groups diagnosed with dementia was possibly due to a bias leading to over-diagnosis of dementia in less intelligent participants or that the NART was sensitive to the dementing process. The small sample size in this study, particularly in regards to the dementia groups ($n = \leq 15$), was a major limitation. As a result of the small numbers in the dementia groups the significant effect of education observed in the initial analysis for predicted IQ was not taken into account when comparing NART predicted IQs for normal, mild dementia and mild/moderate dementia groups for each age group. As age was not observed to contribute to the variance in predicted IQ it would have been preferable to combine the two age groups and split into education groups (age of leaving school groups: $\leq$ 14 years old; > 14 years old).

O'Carroll and Gilleard (1986) proposed that another method of evaluating NART sensitivity to dementia would be to assess the relationship between increasing severity of dementia and impaired oral reading. They studied a group of 30 psycho-geriatric day patients with dementia (age range 65-86 years). They found no significant relationship between measures of severity such as the Clifton Assessment Procedures for the Elderly (CAPE) Survey and the Kew Cognitive Map with NART or the Mill Hill Vocabulary Scale synonym section (MHVS). The sample was divided into diagnostic groups and once again no significant differences emerged on any of the variables. The total sample was then divided into two groups of 15, based on the Information/Orientation scale of the CAPE survey cut off score of ≥ 5 for mild dementia and ≤ 4 for moderate dementia. No significant between-group differences were noted for NART or MHVS. The sample size was extremely small and there was not a broad enough range of dementia severity within the sample to adequately compare NART performance across severity groups. No details were provided regarding the mean scores for the severity measures, which make it difficult to assess the clinical significance of the results. The I/O scale of the CAPE survey may not be a sensitive measure of cut offs for severity of dementia.

4.6 Validity of the short NART

A shortened version of the NART has been developed by Beardsall and Brayne (1990) as they proposed that the list of 50 words can be quite difficult for many patients and can cause distress. Nelson (1982) recommended a cut-off of 14 failures out of 15 consecutive items. However, Beardsall and Brayne suggested that in practice this may mean that the patient actually completes the test, despite difficulty, as the words may not be in increasing order of difficulty for the particular population studied.

Beardsall and Brayne (1990) studied an elderly female rural community sample (n = 316) aged between 70-80 years and provided equations to predict performance on the second half of the test from performance on the first half of the test. They developed a regression equation, to predict scores on the second half of the NART from the score on the first half, from the scores of 122 unimpaired participants and subsequently tested this on another 122 unimpaired, 20 demented and 52 depressed participants. The actual score on the first half of the test was then added to the predicted score for the second half of the NART to produce a total predicted NART score. The total predicted score was highly correlated with the actual total NART (r = .93). It was also significantly correlated in a group of patients with dementia (r = .95) and a group of depressed patients (r = .93). The results showed that persons who score <12 correct are unlikely to correctly pronounce any words on the second half of the NART and the short NART score can be taken as the total. Conversely, those who scored > 20 on the first half were unpredictable for performance on the second half. Therefore the authors suggest those who score > 20 on the first half should complete the second half. It would appear that this would be justified as the patient who exhibits good performance on the first half is unlikely to suffer distress by reading the second half. For those who score between 12 and 20 on the first half of the NART the authors suggest the prediction equation can be used to substitute the score on the second half of the test with a good degree of accuracy.

Crawford, Parker, Allan, Jack and Morrison (1991) examined the accuracy of the Short NART to predict full-length NART with a larger, heterogeneous sample more representative of the general population than the sample studied by Beardsall and Brayne (1990). The sample consisted of 674 participants (F = 378; M = 296) with an age range of 16 - 90 years and representations of each socio-economic group, similar to that represented in the general United Kingdom (UK) population. Furthermore, the authors administered a full WAIS to a sub-sample of 142 participants (F= 67; M = 75) to examine the ability of the short NART to predict WAIS

FSIQ. A correlation of 0.88 between predicted and obtained NART scores was observed for participants for whom the short NART score was taken as the total ($n = 55$) and a correlation of 0.83 for participants whose total score was predicted from short NART errors ($n = 282$). NART predicted IQ or short NART predicted IQ did not differ significantly from the obtained IQs. Both methods of predicting IQ yielded high correlations with WAIS IQs, with a correlation of 0.81 for NART prediction and 0.79.for the short NART prediction. However due to the high correlation between NART predicted IQ and short NART predicted IQ ($r = .99$) it was necessary to examine if the size of the correlation coefficients differed significantly. They found that while the difference in size of these correlation coefficients appear small they were statistically significant ($t = 1.82$, $p = < .05$, one tailed). The authors noted that the distribution of errors in predicting IQ was similar for both methods, although there was a small percentage of participants who exhibited extreme discrepancy scores with the short NART that were not produced with the full NART.

The authors concluded that the accuracy of the short NART in predicting FSIQ is virtually equivalent to the full NART in an unimpaired population. Therefore the short NART can be used in clinical practice if necessary but they suggest that the short NART has limited practical utility. The full NART is quick and easy to administer and score and generally does not cause distress. The susceptibility of clerical error in converting short NART correct scores to NART error scores before obtaining estimating premorbid IQ is a limitation of the short NART (Crawford et al., 1991).

In addition, Bucks, Scott, Pearsall and Ashworth (1996) assessed the utility of the short NART in a clinical sample of patients (*N* = 202) aged between 44 and 88 years, from a UK Memory Disorders Clinic. They concluded that the discrepancies between short NART and full NART error scores were "outside the bounds of clinical and statistical acceptability" (p. 133) for the group studied and suggested the short NART should not be used in clinical practice. The authors found that that short NART underestimated the full NART predicted IQ by 4 – 6 points in 41% of participants scoring < 12 correct (*n* = 17) and 34% of participants scoring between 12 – 20 correct on the short NART (*n* = 20). When comparing the ability of the short NART and the full NART to predict actual obtained WAIS IQ, both methods revealed similar discrepancies with 52.5% of participants having a discrepancy of 0 - 5 IQ points for the short NART and 54.7% having a discrepancy of 0 - 5 IQ points for the full NART. However, the sample in the Crawford et al. (1991) study was much larger for each group than the Bucks et al. study (1996) and the Crawford et al. (1991) study administered the WAIS and was therefore able to compare both short NART and full NART prediction with actual obtained WAIS IQ providing a superior comparison of ability to predict IQ. In comparison, the Bucks et al. (1996) study did not administer a WAIS and compared short NART predicted IQ with Full NART predicted IQ. The discrepancy in IQ points observed in Crawford et al. (1991) study cannot be compared with the Bucks et al. (1996) results as the Crawford study had an obtained WAIS FSIQ whereas the Bucks study was comparing estimated IQs of two prediction methods, suggesting the study by Crawford et al. is more authoritative. Finally, as Beardsall and Brayne (1990) noted, the reliability of the equations should be checked on different samples. Bucks et al. (1996) indicated they followed the procedure outlined by Beardsall and Brayne but they did not provide details of checking the equation on the sample used in their study.

In summary limited research on the validity of the short NART in clinical populations and findings from the studies reviewed suggest the full NART should be used in preference to the short NART. However using the score on the first half as the total NART score if a person scores less the 12 correct appears to be a valid measure of estimating premorbid IQ, in such cases where proceeding would cause distress (Beardsall & Brayne, 1990; Crawford et al., 1991).

4.7 Combined NART-demographic regression equations

Several studies have examined regression equations based on combined NART and demographic variables to estimate premorbid IQ (Crawford et al., 1989d; Crawford et al., 1990c; Crawford et al., 1990d; Berry et al., 1994; Blair & Spreen, 1989; Willshire et al., 1991). These studies were based on the assumption that demographic variables may mediate the relationship between NART performance and IQ. Crawford et al. (1989d) studied a normal UK population ($N = 151$) and found that equations incorporating NART and demographic variables (age, gender and occupation/social class) accounted for 73%, 78% and 39% of the variance for WAIS FSIQ, VIQ and PIQ respectively. Education did not significantly increase the amount of variance, although detailed information about education level of this sample was not provided. The authors suggest that there was no additive effect of education on the proportion of IQ variance predicted due to the amount of covariance between NART and education. NART alone accounted for 66%, 72% and 33% of the variance and demographic variables alone accounted for 50%, 50% and 30% of the variance. The results showed an improvement in prediction accuracy using combined NART and demographic variables.

To determine if this finding would hold in another sample, Crawford et al. (1990c) combined NART and demographic variables in a multiple regression equation using the original NART standardisation sample (N = 120). They also found that combined NART, age, gender and occupation regression equations significantly increased the amount of variance accounted for in WAIS FSIQ, VIQ and PIQ compared to equations based on NART alone. These equations were then applied to a second sample taken from the Crawford et al. (1989d) study and as shrinkage in predicted variance did not occur the two samples were combined resulting in 63%, 66% and 38% of FSIQ, VIQ and PIQ variance accounted for by combined NART/demographic variables. However, it was notable that the Crawford et al. (1989c) study used the full-length WAIS and accounted for more variance (73%, 78% and 39%) than the combined sample in the Crawford et al. (1990c) study using the short-form WAIS (63%, 66% and 38%). Further to this, Crawford et al. (1990d) provided evidence of construct validity for the combined NART and demographic premorbid IQ estimates. They performed a factor analysis on WAIS subtests and the combined NART/demographic estimate (NDE) and found that the NDE loaded highly on the first un-rotated principal component 'g' (0.90). This was higher than any of the WAIS subtests and suggests that the NDE has very high construct validity and is superior to NART alone.

Blair and Spreen (1989) studied the utility of combining demographic variables with their adaptation of the NART for North American participants, the New Adult Reading Test (NART-R). They found that adding demographic variables did not significantly increase the percentage of explained variance in predicted FSIQ in this sample (N = 66). These findings were supported by the results of a UK study by Bright, Jaldow and Koleman (2002). They also found that combining NART error scores with demographic variables did not significantly increase the amount of variance in IQ explained by NART only in either controls (*N* = 51) or a mixed neurological sample of patients (*N* = 98). The limited variability in educational background and occupational status for both samples studied may have reduced the predictive accuracy of demographic variables (Blair &Spreen, 1989).

Conversely, Berry et al. (1994) examined the New Adult Reading Test (NART-R) and demographic equations to predict IQ obtained 3.5 years earlier in a normal elderly US population (*n* =54*)*. They found that NART-R, education and age were important predictors of WAIS-R FSIQ. NART-R, education and gender were significant predictors for VIQ and NART-R and age were significant predictors for PIQ. The authors concluded that the addition of demographic variables to NART-R raised the predictive power of the equation, possibly due to capitalization on variance specific to the sample studied.

Willshire et al. (1991) also examined the addition of demographic variables to NART when predicting IQ in an Australian neurologically normal sample (*N* = 104) and clinically relevant sample (N = 49). The authors noted the later sample was referred to as clinically relevant as they were >55 years of age and only a small proportion of people under 55 years are known to suffer dementia. The authors found that the inclusion of education with the NART provided a substantially higher estimate of premorbid cognitive functioning than NART alone. The combined equations of NART error score and education accounted for 46% of the variance in WAIS-R IQ in the total sample compared to only 26% of the variance in WAIS-R IQ accounted for by NART error score alone. The authors also examined participants < 55years old ($N = 76$) and found that combined NART error score and education accounted for 34% of the variance in WAIS-R IQ, although in the group aged > 55years ($n = 28$) the combined equation accounted for 67% of the variance. Therefore the combined education/NART equation accounted for more FSIQ variance in the > 55years group than in the < 55years group or in the combined sample.

However these equations accounted for substantially less variance than that found by Crawford et al. (1989c) who found 73%, 78% and 39% for FSIQ, VIQ and PIQ respectively using combined demographic/NART error score equations. Notably Crawford et al. (1989c) used the full-length WAIS in comparison to Willshire et al. (1991) who implemented a short version of the WAIS-R. Furthermore, the use of the original NART (Nelson, 1982) to predict WAIS-R IQs (Wechsler, 1981) is a limitation of this study (Nelson & Willison, 1991). Willshire et al. (1991) noted that the WAIS-R scores are not directly comparable to WAIS scores as approximately 20% of items had been replaced in the 1981 revision. The order of subtest administration had also changed and the complete re-norming of the scale in 1981 suggest the WASI and WAIS-R are not directly comparable. Also, the authors did not attempt to assess for dementia in this group of individuals. Considering the heightened risk of dementia in the older age group it would have been preferable to assess participants for dementia using a commonly used test such as the (MMSE) which has been used extensively as a screen for dementia in research and clinical practice (Berry et al., 1994; Crawford et al., 2001; Schmand, Geerlings, Jonker & Lindeboom, 1998; Taylor et al., 1996).

While the NART/demographic method appears promising, accounting for more variance in FSIQ than purely demographic equations, large variability in findings have been observed. This may be due to variation in the composition of the specific samples studied, with varied age ranges and years of education across studies. Further research with large samples of normal and neurologically impaired individuals is required.

4.8 NART adaptations

The British NART (Nelson & Willison, 1991) has had several adaptations and reliability and validity studies provide some support for their use in other populations. A Dutch version of the NART, known as the DART (Schmand et al., 1998) has been published in the Netherlands. Adaptations for the NART to be used in the United States (North American Adult Reading Test [NAART], Blair & Spreen, 1989), Canada (American NART [AMNART], Grober & Sliwinski, 1991) and an Australian version, the AUSNART (Hennessy & Mackenzie, 1995) have been developed. However to date there are limited studies regarding the validity and reliability of these adaptations. As noted above the study by Willshire et al., in which the NART was used with an Australian population had some serious flaws in the design particularly in regards to the use of the original NART to estimate WAIS-R IQs and to the failure to screen for dementia in the older age group.

4.9 The Australian National Adult Reading Test (AUSNART)

The AUSNART was developed by Hennessy and Mackenzie (1994) to address validity issues of administering the NART, developed with British participants, to an Australian sample. The authors suggested that NART should not be used in an Australian population as the NART was based on a prorated seven sub-test short form WAIS-R, which is regarded as inadvisable. They also noted that the pronunciation rules differ between English and Australian populations for a number of words (e.g., *capon, catacomb, bouquet, banal, idyll*) on the NART which may lead to an underestimation of premorbid IQ in an Australian population. The inter-rater reliability for some NART words such as *aeon, puerperal, aver, sidereal and prelate,* have been questioned (Crawford et al., 1989a; Alcott et al., 1999).

The AUSNART was developed with 49 (M = 14, F = 35) first year University of Tasmania psychology students with an age range of 17 to 40 years. The full WAIS-R, NART and an additional reading list consisting of 213 words were administered. The words for the additional reading list were obtained from the Australian Pocket Oxford Dictionary (Turner, 1984) and were selected based on the following criteria: irregular pronunciation; less than 12 letters to minimise adverse effects of stimulus complexity; and European language words which were in everyday Australian usage. The pronunciation guide was based on the Macquarie Dictionary (Delbridge, 1985). Item analysis of both the NART words and the additional word list were performed and a list of 64 words ranging in difficulty from .02 to .96, with scores approaching 1 indicating easier items, were chosen to make up the final AUSNART list.

Hennessey and Mackenzie (1994) found the AUSNART to be a valid and reliable estimate of premorbid IQ revealing a coefficient alpha of .94. The correlations between the AUSNART and WAIS-R FSIQ, VIQ and PIQ were all significant, yielding 0.76, 0.74 and 0.61 respectively, resulting in AUSNART predicting 58%, 55% and 37% of the variance in WAIS-R FSIQ, VIQ and PIQ. In comparison, while still significant, the NART predicted less IQ variance, yielding 33%, 34% and 18% of the variance for WAIS-R FSIQ, VIQ and PIQ in this sample. Hennessey and Mackenzie noted that the results of the NART in their study were similar to that found by Willshire et al. (1991) who also examined NART performance in an Australian sample, although the NART only predicted 26% of the variance in WAIS-R IQ in the Willshire study. These results question the use of NART in an Australian population and support the development of a reading test specifically developed in Australia for use with an Australian population. The AUSNART was shown to be a better estimate of WAIS-R IQ in this Australian sample.

Split-half reliability was found to be .94 for the AUSNART showing a slight improvement on previously reported split-half reliabilities for the NART of .90 and .93 (Nelson, 1982; Crawford et al., 1988b). Both NART and AUSNART correlated highly with general intelligence measured by WAIS-R. As Hennessey and Mackenzie (1994) noted there is a need for further research using the AUSNART as the results of their study may have been inflated due to the same participants being used for both test construction and validation. The study did not examine the effects of demographic variables on AUSNART performance or the addition of demographic variables to the AUSNART to improve estimation of IQ and these are areas where future research is required.

Only two studies have examined the AUSNART in an Australian sample and these have both examined a contextual version, developed specifically for the two studies described. Lucas, Carstairs and Shores (2003) examined the use of a contextual version of the NART for an Australian sample ($n = 244$) aged 18 – 34 years who had taken part in the Macquarie

University Neuropsychological Normative Study (MUNNS). They placed the AUSNART words in sentences in a similar manner as the Cambridge Contextual Reading Test (CCRT), providing context for the words. The authors compared seven methods of estimating premorbid IQ. In regards to reading tests the spot-the-word test and the contextual AUSNART (C-AUSNART) were favoured over the WRAT-3 with the C-AUSNART predicting 35% of the variance for FSIQ and 41% of the variance in VIQ. However, they cautioned against the use of the C-AUSNART with populations of non-English speaking background as it was found to be moderated by language background, although this was a small effect accounting for 1.1% of the variance.

Further to this Carstairs, Myors, Shores and Fogarty (2006) studied the effect of language background of healthy participants on a range of cognitive tests including the C-AUSNART. The sample consisted of participants from the MUNNS study (n = 116) and they found that people with non-English background (first language other than English) had significantly more C-AUSNART errors than those of English speaking background and those of non-English speaking background with English as a first language. These results provide support for the previous findings.

4.10 NART limitations

While the NART is used extensively in neuropsychological practice several limitations need to be considered. The NART cannot be used when a patient has dyslexia, visual impairment or articulatory problems and due to the degree of difficulty of NART words, repeated failures may be discouraging to patients. NART may also underestimate the IQ of self-educated individuals who may be familiar with a low frequency word but not its pronunciation. The NART words cannot be adapted to other languages with a more regular mapping of pronunciation on to orthography due to the orthographic irregularity of the words (Baddelely et al., 1993). In such cases alternative estimates of premorbid IQ as outlined in Chapter 3 need to be considered.

4.10 Summary

This Chapter has reviewed studies which indicate NART is a valid and reliable estimate of IQ. It has been shown to posses good reliability, with split-half reliabilities ranging between 0.90 - 0.93 and test-retest reliability of 0.98. The NART has been found to correlate highly with general intelligence ($r = .85$) and to be more resistant to dementia than WAIS Vocabulary. The irregular words used in the NART have been found to be more resistant to cognitive decline than regular words, suggesting it taps previous knowledge and not current ability. The NART has been shown to account for up to 66% of WAIS FSIQ and 72% of WAIS VIQ in normal British populations and the addition of demographic variables has increased the amount of variance accounted for to 73% and 78% respectively. The results suggest NART is a valid estimate of premorbid IQ in its country of origin. However, NART adaptations in other countries have found NART does not account for as much variance in FSIQ as for British samples. The AUSNART (Hennessey & Mackenzie, 1994) was developed to address concerns regarding the use of the British NART in Australia and Hennessey and Mackenzie found the NART only accounted for 33% and 34% of WAIS-R FSIQ and VIQ and the AUSNART accounted for 58% and 55% respectively. However the AUSNART is limited by the small sample of University psychology students used in the norming study and requires further research. Validation studies of NART with clinical groups reviewed in this Chapter have suggested that NART is a valid estimate of premorbid IQ in clinical samples but these studies have been limited by small and mixed clinical samples and ill-defined assessment time post injury. Further research is required with large well-defined clinical samples and longitudinal designs to enable a more thorough examination of NART sensitivity to cognitive decline and brain injury.

This Chapter also examined the short version of the NART as some patients find the NART extremely difficult. Beardsall and Brayne (1990) examined the use of the short NART and concluded that if a person scores less than 12 on the first half, and proceeding would cause distress then using this score as the total appears valid. However, further studies with large normative and clinical samples are required before this can be used with confidence. Also, limitations of the NART were discussed and in situations where NART is not appropriate, alternate methods for predicting premorbid IQ will need to be considered (see Chapter 3). While the studies reviewed provide some evidence for NART as a valid estimate of premorbid IQ in dementia and TBI, recent research has questioned this and these studies will be reviewed in detail in Chapters 5 and 6.

CHAPTER 5

Can NART Performance Be Impaired?

As reviewed in Chapter 4 studies confirming the validity of the NART, particularly with clinical groups, have consisted of small sample sizes, mixed diagnoses/severity groups and limited age ranges. Few studies have administered the WAIS or WAIS-R, for both controls and clinical groups to compare the predicted versus obtained IQ. Several studies have examined the validity of the NART retrospectively (Berry et al. 1994; Crawford et al., 2001) but few have examined changes in NART score prospectively in a longitudinal design, following traumatic brain injury.

To examine the validity of the NART as a predictor of premorbid IQ in dementia and to address the limitation of previous studies Stebbins, Wilson, Gilley, Bernard and Fox (1990) studied a control group (*N* = 26) and a patient sample (*N* = 199), divided into Very Mild (MMSE: 24 - 27), Mild (MMSE: 16 - 23) and Moderate/Severe (MMSE: 5 - 15) dementia groups based on performance on the MMSE. All patients met criteria for dementia using several standardised scales. The results showed that the very mild dementia patients did not significantly differ from controls in NART performance however all other groups differed significantly from each other on NART performance, with increasing severity of dementia increasing differences between controls were observed. The authors concluded that even in mild dementia NART underestimates premorbid IQ and in moderate/TBI dementia NART grossly underestimates IQ. A significant two-way interaction for dementia severity and educational level suggested that NART scores in the mild group tended to hold for highly educated patients suggesting that education may moderate the impact of dementia on NART performance. There was a significant difference in educational level for the moderate/severe dementia group (*M* =12.4 years) compared to the control group (*M* = 13.3 years), although notably the means are high for both groups. The authors caution the use of NART for patients with MMSE scores in the impaired range (< 24). It has been noted that the MMSE can be affected by level of education in community samples suggesting that some individuals may be misclassified as suffering from cognitive decline when lower scores may be indicative of lower education (Russell et al., 2000).

A study by Patterson, Graham and Hodges (1994) found that NART performance was significantly correlated with performance on semantic memory tasks in patients with dementia of the Alzheimer's type (DAT). The study examined performance on NART and a range of semantic memory tests of varying degrees of difficulty for a control group (*N* = 25) and a group of patients with diagnosis of probable DAT (*N* = 45) divided into severity groups based on MMSE scores. The results indicated that there was a reliable decrement in semantic memory as a function of disease severity. They also found DAT patients had a substantial impairment in word reading of lower frequency exception words and non-words. The NART showed a dramatic increase in errors as a function of severity with a significant difference in NART performance between all groups except between the mild and moderate dementia patient groups. The authors noted that when using the equation provided by Nelson and O'Connell (1978) the NART underestimated premorbid IQ by approximately one standard deviation (15 IQ points) in patients with only moderate dementia. These results provided important information regarding the validity of the NART with patients with dementia;

however, once again the small sample size is a limitation of this study suggesting a need for further research with larger sample sizes.

Fromm et al. (1991) also demonstrated that a modified version of the NART was sensitive to Dementia. They examined a group of patients diagnosed with probable Alzheimer's disease (AD) (N = 18) and elderly controls (N = 20). They collected longitudinal data on NART performance and other neuropsychological test performance over 2 years. The AD group scored significantly more poorly on the NART over time and the controls performed significantly better than the AD group at each test date. A significant correlation between NART and MMSE correlations was observed at the final assessment only. The authors concluded that there was a strong relationship between severity of dementia and NART performance. These results support the findings of Stebbins et al. (1990) who place doubt on the appropriateness of using NART as a predictor of premorbid IQ in moderate and severe dementia. However, the sample size was very small in this study also and the NART modification had not been validated on other unimpaired or clinical populations.

A study by Paolo, Troster, Ryan and Koller (1997) also found NART performance to be impaired in more severely demented patients. They compared NART performance, Barona's (1984) demographic estimated IQs equation and current WAIS-R performance of 44 probable Alzheimer's (AD) patients with 44 matched controls. The AD patients performed significantly more poorly on WAIS-R (M =84.82) than controls (M = 113.02) and had significantly more NART errors (M = 24.75) than the control group (M = 19.14). There were no significant differences in the Barona demographic estimated IQs as the controls were matched demographically to the AD group. A significant difference between obtained FSIQ (*M* = 84.82) and NART predicted FSIQ (*M* = 103.66) for the AD group was observed. Examination of NART predicted FSIQ and VIQ for AD severity groups (Mild & Mild/Moderate) based on the Dementia Rating Scale indicated that all comparisons were significantly different with the moderate group performing more poorly than the mild group. A comparison for NART error scores revealed a trend towards significance with the Moderate group yielding more errors than the Mild group. Paolo et al. note that the failure to reach significance may be due to small sample size and severe Bonferroni correction due to the number of comparisons. The authors concluded that the results indicated NART is sensitive to dementia and that the difference in NART performance for mild and mild/moderate groups is of clinical significance, possibly suggesting intellectual deterioration.

Taylor et al. (1996) examined the validity and stability of the AMNART to estimate premorbid VIQ for patients with DAT (*N* = 40) compared with demographically matched controls (*N* = 40) in a longitudinal study. Assessment of MMSE, Disability Rating Scale (DRS) and AMNART were performed over three years at yearly intervals. The results showed that MMSE score, DRS and AMNART performance, and the corresponding VIQ, significantly declined over time with increasing severity of dementia but remained stable for the matched controls. Taylor et al. produced a correction formula based on performance on the MMSE, an independent measure of global cognitive impairment, to account for the negative effects of increasing dementia severity on AMNART performance. They developed the correction formula on 20 DAT patients with the highest AMNART estimated VIQ scores at the first assessment and then validated the formula on the remaining 20 DAT patients. The correction formula is shown below:

[corrected premorbid VIQ] = 118.2 - .89 [AMNART errors] + .64 [years of education] + 0.63 [30-MMSE score].

When the correction factor developed by Taylor et al. (1996) was applied to the DAT patients' data the resulting estimated VIQ did not significantly differ from the demographically matched controls at any of the three assessments. The correction formula enables a more accurate estimate of premorbid VIQ in patients with DAT as the validity of AMNART, and NART, has been shown to be notably reduced as the disease progresses (Fromm et al., 1991; Stebbins et al., 1990; Patterson et al., 1994). However the formula in this study was developed on a small sample and with patients with high level of education (M = 14.4 years) and further studies validating this formula are required.

An important study by Schmand et al. (1998) examined the decline of IQ estimates using the Dutch version of the NART (DART) as a function of dementia severity and also provided a correction formula based on MMSE performance. The authors provided unique data on the effect incident dementia may have on reading ability, retesting a large group of elderly people consisting of a 'suspect' Dementia group (N = 197) and a control group (N = 117) after a six-year period. Patients were considered 'suspect' for dementia if they had scored $\geq$ 24 on MMSE at the initial testing but $\leq$ 23 at the 6 year follow-up, producing a mean MMSE score of 20. The healthy controls produced a mean MMSE score of 27.5. The 'suspect' group was then divided into five dementia severity groups, ranging from no dementia to severe dementia. The results indicated that the DART remained stable over the 6 year period for healthy elderly individuals. In the clinical group, those who were 'suspect' but without dementia, the mild and minimal dementia patients showed worsening DART performance at a rate of 1 IQ point per 2 years on average. However, the DART declined considerably in moderate and severe cases of dementia over the 6 year period. Patients with moderate dementia (n = 17) revealed a decrease of 14.9 IQ points over the 6 year period and a decrease of 18.5 IQ points was observed for the severe dementia patients (n = 2) over the same period. Notably, the patient numbers in these groups were significantly smaller than the other groups. To correct the observed IQ of patients with dementia on the basis of their MMSE score the authors used the following formula:

$$\text{Corrected IQ} = \text{observed IQ} + 0.045\ \text{MMSE}^2 - 2.44\ \text{MMSE} + 35.7$$

When the authors applied this formula to the follow-up IQs of the suspect group, the decline was eliminated completely. However as the formula was derived from the same data it would be expected that the decline would be eliminated. It is therefore, essential that this formula be validated with other samples.

Schmand et al. (1998) also noted that verbal abstraction and category fluency, both components of semantic memory, best explained decline in DART IQ, providing support for the findings of Patterson et al. (1994) in suggesting that the decline of NART IQ scores is ascribed to a breakdown in semantic memory. It was concluded by Schmand et al. (1998) that in mild or questionable cases of dementia reading ability can still be considered a valid estimator of verbal intelligence. In cases of moderate dementia the scores on reading tests underestimate the premorbid level by approximately one standard deviation (Schmand et al.).

While the implementation of a longitudinal design provided best evidence compared to the commonly used cross-sectional design in studies with Alzheimer's patients, Schmand et al. (1998) noted several limitations of their study. The correction method for IQ as a function of MMSE should be used with caution because it involves some circularity as the MMSE correlates moderately with education and intelligence even in cognitively intact people. Therefore the correction method should not be used when the MMSE score is not abnormal. There are issues with the psychometric weaknesses and omnibus features (a number of short strongly varying tasks that result in a single sum score) of the MMSE. It yields only modest reliability and its core distribution is skewed and measures a global cognitive deterioration only. The authors suggest that it may be wiser to correct on the basis of language decline; however this also poses the problem of circularity as language functions are inherently related to verbal intelligence. Overall the authors suggest that the correction method for the DART is still useful in the estimation of premorbid IQ for patients with dementia, however further research is required.

5.1 Summary

The studies reviewed have provided a thorough examination of the degree of NART sensitivity to dementia. These studies have increased knowledge in the field, providing strong evidence for NART sensitivity to cognitive decline according to dementia severity. Schmand et al. (1998) and Taylor et al. (1996) examined versions of the NART, a Dutch (DART) and an America (AMNART) version, over a six year and three year period respectively, and both were shown to be sensitive to cognitive decline. The authors also provided clinicians with equations to correct the underestimation of premorbid IQ in patients with dementia. Other studies reviewed compared controls with patients of different dementia severity and found those with severe dementia show poorer NART performance than controls.

Together these findings suggest NART is sensitive to cognitive decline and the degree of sensitivity is related to the severity of dementia, with greater severity showing poorer NART performance. This raises the question as to whether the NART is also sensitive to brain injury. If the NART is sensitive to brain injury, then questions arise as to whether it is also affected by the severity of injury and whether other factors may also influence NART performance in TBI. These questions will be examined in Chapter 6.

CHAPTER 6

NART Performance in Traumatic Brain Injury

The significant long term deficits experienced following TBI and the evaluation of rehabilitation alternatives necessitate the assessment of premorbid intellectual ability. Knowledge of a person's premorbid ability is not often readily available through prior testing and alternative measures are required to estimate premorbid IQ. While the NART has been shown to be a valid estimator of premorbid IQ in some clinical groups (Crawford et al.1988; Crawford et al.,1987; Crawford, 1989; Nelson & O'Connell, 1978; and O'Carroll, Baikie & Whittick, 1987), the assumption that NART performance is totally insensitive to brain damage has been challenged in recent research and an increasing body of evidence suggests that the NART may underestimate premorbid IQ when used in TBI (Freeman & Godfrey, 2000; Freeman, Godfrey, Harris & Partridge, 2001; Riley & Simmonds, 2003; Skilbeck, Allen & Brechin, 2005).

Earlier studies such as those conducted by Crawford et al. (1988), finding no differences in NART scores between TBI patients and matched controls, have some limitations which should be considered when interpreting the results. Firstly, the Crawford et al. study was relatively small consisting of 18 TBI patients and secondly, it revealed only modest power indicating it was unlikely to have detected a genuine medium-sized effect (Riley & Simmonds, 2003). No details regarding severity of injury or assessment time post-injury were given: these variables are necessary for adequate interpretation of results and should be considered in future research.

A more recent study by Watt and O'Carroll (1999) also observed no differences between TBI patients ($n = 25$), orthopaedic trauma patients ($n = 20$) and healthy controls ($n = 50$) on NART error scores. Notably, this study showed only modest statistical power and the TBI sample consisted of mixed severity groups, which may have resulted in dilution of effects. The mean time post-injury was 23.8 months post-injury resulting in assessments conducted at various stages of TBI recovery and hence probable recovered NART performance at time of testing, particularly for mild TBI patients.

Crawford et al. (1990a) developed a regression equation to address issues of concern regarding NART impairment with clinical populations. A sample of 659 healthy participants was used to build a regression equation using demographic variables (years of education, socio-economic status, age and gender) to predict the NART score. The multiple correlation between demographic variables and NART score was highly significant (0.70). Crawford et al. concluded that an obtained NART error score ≥ 11.4 points higher than predicted indicated NART performance was impaired. The equation has not been used extensively in research, although findings from a study by Skilbeck et al. (2005) supported its use as a valid estimate of NART performance in a UK neurological sample ($N = 175$).

To examine possible NART impairment, Freeman et al. (2001) examined the utility of the demographic equation developed by Crawford et al. (1990a) in detecting impaired NART performance in patients with TBI in a New Zealand sample. Freeman et al. applied the equation to three samples, a community group ($n = 80$), a traumatic brain injured group ($n = 65$) and an orthopaedic control group ($n = 27$). The results showed that 30% of the TBI sample had impaired NART performance, although 18% of the orthopaedic group had discrepant NART scores. However, this was still a significantly lower percentage than the TBI group. The authors suggested the 18% figure may reflect an increased risk of undiagnosed brain injury in the orthopaedic group or other non-specific factors associated with the illness. In comparison the community group had only 11% of participants showing significantly discrepant scores. Therefore, NART scores for one third of the TBI patients would have been impaired and would have resulted in an underestimation of the cognitive decline experienced by the patient. The authors concluded that the equation should be used by clinicians to check the accuracy of NART performance before using it to estimate premorbid IQ. Details of severity of injury and time post-injury for the TBI group were not given in this study and therefore affect the applicability of the findings. Given the Crawford et al (1990a) demographic equation was developed on a normal sample it is not clear how accurate in will be in TBI samples (Riley & Simmonds, 2003). TBI populations generally have a higher rate of premorbid difficulties, including poor academic performance and substance abuse, resulting in a lower premorbid IQ than their demographically-equivalent counterparts. Riley and Simmonds suggested the equation may overestimate premorbid IQ in the TBI sample rather than reflect impairment of NART performance. Conversely, it is possible that the NART estimate equation, based on demographic information such as education and occupation, would reflect poor academic performance and occupation level, which is thought to be representative of this population.

Further to this, Riley and Simmonds (2003) compared obtained NART scores and estimated NART scores of 26 participants with TBI, using the equation developed by Crawford et al. (1990). The results showed that 30.8% of participants yielded an obtained NART error score significantly higher than the estimated NART error score at their first assessment, confirming the results of the study by Freeman et al. (2001), which observed a similar rate of 30% in a TBI sample. Analysis of the second NART assessment revealed only 7.7% (two participants) of the sample had obtained scores significantly higher than their estimated NART error score. This compares well with the 11% reported for the healthy control group examined in the Freeman et al. (2001) study, suggesting that recovery in processes required to read the NART had occurred by the second assessment.

A study by Skilbeck et al.(2003) examined the applicability of Crawford et al.'s (1990a) demographic equation in the prediction of NART performance in a mixed neurological sample (N=175) and also examined whether NART was impaired in this population. Based on the development of the demographic equation by Crawford et al., it was proposed that a NART error score of 11.4 points above that predicted by the demographic equation indicated an underestimate, reflecting an "impaired" NART. Skilbeck et al. (2003) found 15 (8%) people in their sample showed impaired NART performance. This sub-group of 15 consisted of five participants with closed head injury (CHI) and five participants had other diffuse cerebral damage. That is, 66% of the NART impaired sub-group had diffuse cerebral damage compared to 45% in the NART non-impaired subgroup. Comparisons of the CHI NART-impaired group (n = 5) and CHI NART non-impaired group (n = 39) revealed no significant differences in age, educational background or PIQ but the NART-impaired CHI group made significantly more NART errors (M = 42.8, SD = 4.0) and yielded a significantly lower VIQ (M = 81.8, SD = 8.4) than the NART non-impaired CHI group who produced fewer errors (M = 23.7, SD = 8.8) and a higher mean VIQ (M = 94.8, SD = 11.9). There was a significant difference between the groups in the size of the discrepancy between NART predicted PIQ and obtained PIQ and while the difference between the two sub-groups was large for VIQ it did not reach significance. When combining the NART-impaired CHI patients with NART-impaired patients with other diffuse cerebral damage (n = 10) and comparing this group to all NART non-impaired patients with diffuse cerebral damage (n = 65) similar results were obtained, with both VIQ and PIQ predicted – observed discrepancies significant. The authors comment that the results indicate that either the NART impaired group was of a lower premorbid VIQ or they had suffered greater diffuse damage than the NART non-impaired group. As there was no significant difference between groups on PIQ it would appear that greater diffuse cerebral damage was not the cause as PIQ is often more sensitive than VIQ to this damage and therefore lower premorbid verbal IQ may have contributed to impaired NART performance.

In summary, the authors suggest that the risk of an impaired NART performance is greater in patients with diffuse brain damage who had lower premorbid VIQs. They also recommend the use of Crawford's equation to predict NART performance and detect impaired NART performance in neurological patients. Crawford et al. suggested the study confirmed the use of the demographic equation for predicting NART score. However the number of low outliers in the general population was 5% and it only reached 8% in this neurological population suggesting that if the NART were impaired in this population then a higher percentage of cases would reveal impaired NART scores. However, as noted, the NART impaired subgroup had a very small number of participants (n = 15) and 66% of these had diffuse cortical damage compared to 45% in the NART unimpaired subgroup. Therefore replication of these results in a larger sample is required.

To further examine NART sensitivity to TBI, Riley and Simmonds (2003) investigated NART performance longitudinally in patients with severe TBI. Twenty-six patients were assessed with NART within the first 12 months post-injury and again 12 months later. The authors also examined demographic estimates of IQ, developed by Crawford et al. (1997), and a vocabulary (WAIS-R subtest) estimate of VIQ administered to16 participants at the same time as the second NART. They found significantly (p = .002) more NART errors made on the first NART (M = 26.3) compared to the second NART (M = 21.8). When converting NART error scores to estimate verbal IQ scores, the first NART estimated VIQ mean was 97.4 (SD = 13.4) compared to the second NART estimated VIQ with a mean of 102.5 (SD = 10.2). Whereas 42% of the sample revealed an improvement of 5+ IQ points, there was a large variation in improvement in estimated IQ over time, ranging from an improvement of 20 IQ points to a decrease of 8 IQ points. Comparisons of the three estimates of IQ revealed that the first NART estimate (M = 97.4, SD = 13.4) was significantly lower (p = .001) than the demographic estimate (M = 104.9, SD = 9.6) and the vocabulary estimate (M = 107.2, SD = 12.1). The second NART estimate (M = 102.5, SD = 10.2) was significantly lower (p =.01) than the vocabulary estimate but comparable to the demographic estimate. The authors suggest the vocabulary estimate may be more accurate than the NART for this sample as there were some individuals whose premorbid IQ was in the 'superior' or 'very superior' WAIS-R IQ range and NART has been shown to underestimate high IQ (Nelson & Willison, 1991).

Riley and Simmonds (2003) concluded that the use of the NART within 12 months of a severe brain injury may substantially underestimate premorbid IQ leading to an underestimation of impairment in cognitive functioning. Nelson and Willison (1991) found that only five percent of healthy adults had a predicted minus obtained verbal IQ discrepancy of 12+ IQ points and suggested this be the cut-off point in diagnosing intellectual impairment. Using this cut-off point in the Riley and Simmonds study, three of the participants would have been diagnosed with intellectual impairment based on initial NART score. However, the second NART would have only diagnosed seven participants as intellectually impaired, suggesting the initial NART underestimated IQ and therefore, underestimated intellectual impairment. The longitudinal design of this study is preferable to the commonly used cross-sectional design in TBI research. The stringent criteria for inclusion of severe TBI only, using specified criteria of a Glascow Coma Scale of 8 or less or PTA in excess of 24 hours, addresses confounding factors of small, mixed severity samples and undefined criteria for measuring severity found in previous studies (Crawford, et al.1988; Watt & O'Carroll, 1999). While the sample size appears relatively small in the Riley and Simmonds (2003) study the authors note that a sample size of 25 will ensure an acceptable probability of detecting a medium effect ($\beta = 0.8$; $\alpha = 0.05$, one tailed). However, the analysis comparing NART predicted IQ with WAIS-R obtained verbal IQ consisted of 16 participants only indicating the sample size was too small to adequately detect effects. There was a large variation in time post-injury for the first NART (1 – 12 months post-injury) and time between first and second NART (12 – 206 months). There are difficulties accessing and assessing patients very early after a severe head injury but it would be preferable to have a defined time post-injury for assessment. Riley and Simmonds also noted that a prospective design rather than the retrospective design utilised would help to clarify reasons for improvement in NART score at the second assessment. They suggest some factors that may influence initial NART impairment and then recovery may be cognitive deficits arising from the TBI, including more general deficits such as visual processing or a deficit more specific to reading. The effect of motivation may also be a contributor to poor performance and could be examined in a prospective design. Watt and O'Carroll (1999) noted that higher depression ratings were associated with increased NART errors scores in their study. They commented that while previous research suggested NART was not affected by depression these studies had examined patients with clinical depression. While it would be expected that if depression was going to affect NART performance then assessing those most vulnerable, those diagnosed with clinical depression, would be more likely to show effects. Alternatively, it may be that the combination of TBI and raised levels of depression may contribute to poorer NART performance. Further studies are required to explore the reasons why impaired NART performance occurs following TBI.

Further evidence of impaired NART performance in TBI patients was provided in a study by Morris et al. (2005). They examined 55 TBI patients of varying severity; on average 7.1 years post-injury. They examined NART performance and compared this with performance on the Cambridge Contextual Reading Test (CCRT), developed by Beardsall and Hupert (1994). Morris et al. showed that both NART and CCRT performance were significantly correlated with GCS ($p = 0.01$, $p = 0.02$ respectively) but not PTA (p=0.07, $p = 0.06$ respectively) and the strength of the relationship was less for the CCRT than the NART, albeit not significantly so. Presence of a coma at any stage was significantly correlated with worse NART and CCRT performance (p= 0.02, $p = 0.04$ respectively). The results suggest that NART is affected by severity of head injury and to a slightly lesser degree, so is CCRT. The authors recommended comparing NART estimates with demographic estimates together with injury severity to assist in identifying instances when NART or CCRT estimates are likely to underestimate true premorbid ability. Morris et al. did not address the possible reasons for the non-significant correlation of NART and CCRT with the PTA measure of severity. It was noted that GCS had a slightly larger range of NART error scores for the *Minor* category to the *Severe* category than the PTA and it is possible this may have contributed to the slightly larger correlation between NART and injury severity reaching significance for the GCS but not for the PTA severity ratings. There were small unequal numbers in each group and the mean time between injury and assessment was 7.1 years, (SD = 0.8) which is a significant amount of time since injury in regards to stage of recovery. To examine NART impairment, assessment of the mild/moderate (based on PTA) or mild and minor (based on GCS) groups would be preferable as close to injury as possible as they would be expected to have improved, possibly to premorbid IQ functioning levels, well before the 7 years post-injury time-point used in this study. This is reflected in the NART error scores for these groups being 20.8 and 19.0 respectively, falling in the top end of average for the corresponding IQs in a healthy population (Nelson & Willison, 1991). This suggests that the sample of milder TBI patients were of high premorbid intelligence and had fully recovered due to the extended period of time since injury. It is notable that even the severe groups in this study yielded NART error scores with corresponding IQs in the average range, albeit the lower end, again suggesting a sample of patients falling in the average to above average ranges premorbidly.

6.1 Summary

The studies reviewed provide evidence for NART sensitivity to TBI, suggesting further research is required to clarify the degree of NART sensitivity to TBI and to examine whether injury severity and demographic variables influence the scores. Whereas studies to date have contributed important information to the field, they have been limited by small samples and have either examined severe TBI only or have used small, mixed severity samples. Riley and Simmonds (2003) proposed that if NART was sensitive to TBI then it would be more likely to be detected in those more vulnerable to the damaging effects of TBI so they examined severe TBI patients only. They provided the only longitudinal study of NART sensitivity in severe TBI, although assessment time post injury varied greatly. Other studies have not utilised the longitudinal design, and have compared controls with TBI samples, examined correlations between measures of severity (GCS and PTA), or examined the Crawford et al. (1990a) demographic regression equation to predict NART performance in TBI samples. While these studies provide important information there is a need for large studies utilising longitudinal within subject designs to provide a more robust assessment of NART sensitivity.

Given the findings of Riley and Simmonds (2003) showing NART sensitivity to severe TBI, questions are raised regarding NART sensitivity in mild and moderate TBI – do people with mild and moderate TBI also experience NART sensitivity, particularly if assessed soon after injury? How long does it take for NART performance to improve and hence represent a more accurate estimate of premorbid IQ? As noted by Riley and Simmonds a more stringent test of the NART's sensitivity would be provided by restricting assessment to the early stages of recovery. Therefore Study 1, presented in Chapter 7, addresses the limitations of previous studies by examining NART sensitivity to TBI in mild, moderate and severe TBI with assessments conducted within one month of TBI and at 6 and 12 months post-injury, using a within-participants design. As NART is a commonly used measure in the estimation of premorbid IQ, often referred to as the 'gold standard', it will be important to clarify the degree of NART sensitivity to TBI. As practice effects have been found to be of little significance in healthy controls any improvement in NART performance over time for TBI patients is likely to represent genuine improvement and suggest NART performance is sensitive to TBI.

CHAPTER 7

Study 1: Examining NART Sensitivity to Traumatic Brain Injury

While early studies indicated the NART was relatively insensitive to traumatic brain injury (TBI) recent research indicates NART may be impaired when used to estimate premorbid IQ for persons with TBI and these studies were reviewed fully in Chapters 5 and 6. Vanderploeg and Schinka (1996) noted all current measures of ability can be adversely affected by some types of brain dysfunction and it is therefore clinically useful to establish the degree of sensitivity of the NART to a very common cause of brain dysfunction, TBI.

In summary, findings from the previous research reviewed in Chapter 6, examining the sensitivity of NART to TBI, have been inconsistent. Small sample sizes, mixed severity samples, ill-defined injury severity and a failure to control for assessment time post injury have contributed to variable findings. Large studies on neurological samples, such as TBI, are required to fully examine possible NART underestimates of premorbid IQ and the effects of injury severity and demographic variables (e.g., age, gender and occupation) on NART performance in these samples.

7.1 Aims and hypotheses

The present study aimed to determine the clinical utility of NART as an estimate of premorbid IQ in an Australian TBI sample, addressing the limitations of previous research. Specifically, the study aimed to examine if NART scores are impaired in TBI if administered soon after injury. If this is so, the study then aimed to determine whether underestimates and recovery are differentially affected by injury severity as noted for other cognitive abilities following TBI (Ponsford et al., 1995). The effects of demographic variables on NART performance were also examined. Based on the research findings reviewed in this thesis the hypotheses were:

1. NART would underestimate premorbid IQ if administered within one month of a TBI and show improvement by 12 months post injury.
2. Participants with severe TBI would show significantly more NART errors than participants with moderate or mild TBI.
3. Participants with severe TBI would show significantly fewer NART errors by12 months post injury compared to the initial assessment.
4. Participants with mild and moderate TBI would show significantly fewer NART errors by six months post injury compared to the initial assessment.
5. Years of education would have a significant effect on NART performance with participants with higher years of education showing fewer NART errors than those with fewer years of education.
6. Socio-economic status (SES) would have a significant effect on NART performance, with participants with higher SES producing fewer NART errors than participants with lower SES.
7. Age would not have a significant effect on NART performance.
8. Gender would not have a significant effect on NART performance.

7.2 Method

7.2.1 Participants

A sample of 194 participants with TBI was obtained from the Neurotrauma Register Research database (NTR). The NTR is a population study for Southern Tasmania, following recovery after TBI for all adult cases who present to the Emergency Department of the Royal Hobart Hospital. It is a joint project of the University of Tasmania and the Royal Hobart Hospital and is funded by the Motor Accident Insurance Board (MAIB). Ethical approval was obtained from The Tasmanian Health and Medical Research Ethics Committee (approval number: H0007116).

Participants were selected on the basis they had completed a NART within 1 month of injury and at 6 and 12 months post injury (+ / - 2 months of the 6 and 12 month assessments). The sample consisted of 117 males and 77 females. Length of time in post-traumatic amnesia (PTA) was the measure of severity as defined by Jennett and Teasdale (1981): Group 1 = Mild PTA $\leq$1 hour (n = 91); Group 2 = Moderate PTA > 1 hour <1 day (n= 53); and Group 3 = Severe PTA $\geq$1 day (n= 50).

Four age bands were used: Group 1 = 16 – 24 years (n = 66); Group 2 = 25 – 40 years (n = 46); Group 3 = 41 – 60 years (n = 54); and Group 4 = 61 – 80 years (n = 28). The age groups were chosen to represent a young group, a young middle aged group, an older middle aged group and an older group. Group 1 represents the age group most commonly reported to experience brain injury, predominantly represented by males, and the older age group is thought to show an increase in the incidence of TBI due to falls (Ponsford et al., 1995).

Occupation based socio-economic status was determined by the Australian Bureau of Statistics Australian Standard Classification of Occupations (Australian Bureau of Statistics [ABS] 1997) system. This system provides nine occupation categories consisting of five skill levels as shown:

1. Managers (1)
2. Professionals (1)
3. Associate Professionals (2)
4. Tradespersons and Related Workers (3)
5. Advanced Clerical and Service Workers (3)
6. Intermediate Clerical, Sales and Service Workers (4)
7. Intermediate Production and Transport Workers (4)
8. Elementary Clerical, Sales and Service Workers (5)
9. Labourers and Related Workers (5).

Fifty percent of the sample could not be coded for SES due to limited information on occupation status with 9% of the sample unemployed, 21% were students and 20% were on a disability pension or retired. A sample of 97 participants with occupation information was coded for SES. To ensure adequate numbers in SES groups some categories were combined: Group 1 = Professionals: categories 1, 2 and 3 (n = 17); Group 2 = Tradespersons: categories 4 and 5 (n = 21); Group 3 = Semi-skilled: categories 6 and 7 (n = 21); and Group 4 = Unskilled: categories 8 and 9 (n = 12).

The sample was split into two education groups based on a median split at 11 years. Group 1 consisted of those who attended school for ≤ 11 years education (n = 87) and Group 2 consisted of those who attended school for > 11 years (n = 107).

7.2.2 Materials

The National Adult Reading Test (NART; Nelson & Willison, 1991) was administered to participants at all three time-points. The NART is a 50-item phonetically irregular word list presented in order of increasing difficulty (Appendix A1). The words cannot be pronounced by common rules of pronunciation such as phonetic decoding. The subject reads the words aloud and error scores were used to estimate WAIS-III FSIQ.

7.2.3 Procedure

Participants were administered the NART word list by trained research assistants at the Neurotrauma Register, as part of a large neuropsychological battery. Training included instruction in administration and pronunciation of NART words and a tape of the words with the correct pronunciation was given to all research assistants (n = 15). Standardised administration, as outlined in the manual, was adhered to. Error scores were calculated by research assistants and all response sheets checked by a Clinical Neuropsychologist and corresponding FSIQ, VIQ and PIQ's from the WAIS-III provisional norms (Crawford, 1997) were recorded (see Appendix A2 for conversion of NART errors to WAIS-III FSIQ).

7.2.4 Design and analyses

A prospective longitudinal, within-participants design was implemented with NART errors as the repeated measure (dependent variable). Three participants who had mean change scores with z-scores above 3.29 were deemed outliers (Field, 2005) and removed from the data set resulting in a total sample of 194 TBI participants. Examination of these participants did not reveal any characteristics that separated them from the total sample. The unusually large fluctuation in NART errors scores observed for these participants suggested they were not characteristic of the population studied and should therefore be removed from the data set (Tabachnick & Fidell, 2001). Results were analysed using SPSS statistical package (SPSS Graduate Pack 15.0 for Windows, 2006). All significance levels were reported at $p = .05$, or better and Bonferroni adjustments made for multiple comparisons.

If NART performance is sensitive to TBI then it is likely that severity of TBI would influence sensitivity, as found in studies for dementia. Therefore, severity was paired with demographic variables (age, education, gender and SES) and three-way mixed ANOVAs with NART error score as the dependent variable (repeated measure) were performed to examine how these variables interact and influence NART performance in TBI. Pairwise comparisons were reported in preference to post-hocs as they take into account unequal sample sizes, hence the means reported in these comparisons are based on estimated marginal means and therefore are slightly different to the means reported in the descriptive statistics. The population studied (TBI) is not likely to have a normal distribution and adjustments would lose valuable information in the data. While ANOVAs assume normality they are thought to be generally robust to violations of this assumption (Field, 2005). Significant ANOVAs were examined with t-tests which provide alternative tests for unequal variances."

As the sample size in this study was large, small differences in group variances were likely to produce a significant Levene's test (assumption of homogeneity of variance), as is common in repeated measures designs (Field, 2005). If the assumption of homogeneity of variance was violated, examination of variance ratios was conducted and if the ratio was less than two it was deemed safe to assume homogeneity of variance (Field, 2005). If variance ratios were above two then a more stringent significance level of $p < .01$ was used (Tabachnick & Fidell, 2001).

Power (β) and effect sizes were reported with a value of 0.8 indicating an 80% probability of detecting a genuine effect. Effect size was measured by Partial Eta Squared ($\acute{\eta}^2$) with .01 a small effect, .06 a medium effect and .14 a large effect (Tabachnick & Fidell, 2001).

NART error scores are reported in the results and the conversion of NART errors to WAIS-III FSIQ, VIQ and PIQ are presented in Appendix A2. Data for NART errors, estimated FSIQ, VIQ and PIQ, and demographic details are presented in Appendix B1.

7.3 Results

7.3.1 Descriptive Statistics

The sample ($N = 194$) was generally representative of a TBI sample as shown in Table 7.1, consisting of 117 males (60%) and 77 females (40 %). Notably the percentage of females is a marginally higher than generally observed for TBI populations (30%). Mean years of education were 12.26 years and mean age was 38.61 years. Mean estimated FSIQ at the initial assessment was 102.65, 104.83 at 6 months and 105.58 at 12 months (see Appendix B2 for frequencies).

Table 7.1

Descriptive statistics for the total sample (N=194).

	Mean	SD	Median	Range
Age (years)	38.61	18.17	35	16-80
Years of education	12.26	2.77	12	6-22
NART err 1 month	20.13	8.92	20	2-47
NART err 6 months	17.97	8.44	17	1-47
NART err 12 months	17.26	8.54	17	2-48
FSIQ 1 month	102.65	9.22	103	75-121
FSIQ 6 months	104.83	8.72	106	75-122
FSIQ 12 months	105.58	8.80	106	74-121
PTA (days)	1.15	3.30	0.05	0 – 30

As shown in Table 7.2 motor vehicle accidents were the most common cause of traumatic brain injury in the sample and falls and assaults were the second and third most common causes of injury.

Table 7.2

Frequencies and percentages of cause of injury.

Cause	*N*	*Percentage*
Motor vehicle accident	69	36
Fall	56	29
Assault	41	21
Sporting accident	14	7
Other	14	7
Total	194	100

7.3.2 Severity & age analyses

A three-way ANOVA (Severity x Age x Time post injury) with NART score as the dependent variable was performed and SPSS print out for analyses are presented in Appendix B3. The total sample (N = 194) was divided into TBI severity groups and age groups as shown in Table 7.3.

Table 7.3

Number of participants in each age group for each severity group.

Severity	16 -24 yrs	25 – 40 yrs	41 – 60 yrs	61 – 80 yrs	Total
Mild	34	20	25	12	91
Moderate	18	12	15	8	53
Severe	14	14	14	8	50
Total	66	46	54	28	194

Time post injury

The results showed a significant main effect of time post injury on NART scores, $F(2, 364) = 22.10$, $p = .000$, partial $\eta^2 = .12$, $\beta = 1$. Mean NART errors at the initial assessment were 19.73 ($SE = 0.68$), 17.91 ($SE = 0.64$) at the 6-month assessment and 16.93 ($SE = 0.65$) at the 12-month assessment. Pairwise comparisons showed a significant decrease in NART errors at the 6-month assessment compared to the initial assessment ($p = .001$) and a clear trend towards a significant decrease in NART errors between 6 and 12 months ($p = .06$). There was an overall significant decrease of 2.80 ($SE = 0.43$) mean errors by the 12-month assessment compared to the initial assessment.

Severity of injury (PTA)

The results showed a significant main effect of severity, $F(2, 182) = 3.24$, $p = .04$, partial $\eta^2 = .03$, $\beta = .61$. Pairwise comparisons showed that overall the Mild TBI group ($M = 16.68$, $SE = .88$) produced significantly fewer NART errors than the Severe TBI group ($M = 20.29$, $SE = 1.13$) ($p = .04$), a difference of 3.62 ($SE = 1.43$) mean errors. To further examine the significant main effect of severity independent samples t-tests were performed and showed no significant group differences in NART performance at the initial or 12 month assessments. However, the Severe group produced significantly more NART errors than the Mild group, $t(139) = -2.72$, $p = .007$ at the 6 month assessment.

Age

There was a significant main effect of age, $F(3, 182) = 4.72$, $p = .003$, partial $\eta^2 = .07$, $\beta = .89$. Pairwise comparisons showed that overall participants aged 16 - 24 years produced significantly more NART errors (M = 21.56, SE = 1.03) than participants aged 41 - 60 years (M = 17.37, SE = 1.09), p = .04 and participants aged 61 – 80 years (M = 15.04, SE = 1.50), p = .005. To further explore the significant main effect of age independent samples t-tests were performed for each assessment and Table 7.4a shows group means and standard deviations and Table 7.4b shows t-values and significance level for each comparison. The results showed that at the initial assessment participants aged between 16 – 24 years produced significantly more mean NART errors than participants who were 25 – 40 years, 41 – 60 years and 61 – 80 years. Participants who were 25 – 40 years produced significantly more mean NART errors than participants aged 61 – 80 years.

Table 7.4a

Means (standard deviations) of NART error scores according to age for all three assessments.

Age Group (*n*)	< 1 month	6-month	12-month
16 – 24 years (55)	23.24 (8.29)	20.03 (7.69)	19.73 (8.05)
25 – 40 years (32)	20.22 (7.47)	17.76 (8.26)	16.96. (7.94)
41 – 60 years (34)	18.41 (9.50)	16.74 (8.58)	16.30 (8.56)
61 – 80 years (20)	15.96 (9.28)	15.14 (9.13)	13.82 (9.31)

Table 7.4b

The t values and significance levels for age group comparisons of NART errors for all three assessments.

	< 1 month		6-month		12-month	
Age Group (*df*)	*t*	*p*	*t*	*p*	*t*	*p*
1 vs 2 (110)	1.98	.05	1.69[a]	.10	1.80	.07
1 vs 3 (118)	2.97	.004	2.42	.02	2.26	.03
1 vs 4 (92)	3.76	.000	2.83	.006	3.10	.003
2 vs 3 (98)	1.07[b]	.29	0.61	.55	0.40	.69
2 vs 4 (72)	2.17	.03	1.27	.21	1.54	.13
3 vs 4 (80)	1.12	.27	0.78	.44	1.21	.23

Note. If Levene's test was significant then the analysis for unequal variances was reported.
[a] df = 92.47. [b] df = 97.42.

At the 6 month and 12 month assessments participants aged between16 – 24 years produced significantly more mean NART errors than participants who were 41 – 60 years and 61 – 80 years. There were no other significant age-group differences in NART error scores at any time point and no significant two-way or three-way interactions

7.3.3 Severity and education analyses

A three-way ANOVA (Severity x Education x Time post injury) with NART score as the dependent variable was performed and SPSS print out for analyses are provided in Appendix B4. The total sample (N = 194) was divided into TBI severity and education groups and group numbers are shown in Table 7.5.

Table 7.5

Number of participants in each education group for each severity group.

Severity	≤ 11 years	>11 years	Total
Mild	35	56	91
Moderate	25	28	53
Severe	27	23	50
Total	87	107	194

Time post injury

There was a significant effect of time post injury on NART errors, $F(2, 376) = 28.15, p = .000$, partial $\eta^2 = .13$, $\beta = 1$. Mean NART errors were 20.55 ($SE = 0.62$) at the initial assessment, 18.57 ($SE = 0.60$) at the 6-month assessment and 17.57 ($SE = 0.62$) at the 12 month assessment. Pairwise comparisons showed a significant decrease in mean NART errors at the 6 month assessment compared to the initial ($p = .000$) and a significant decrease in mean NART errors at the 12-month assessment compared with the 6-month assessment ($p = .04$). There was an overall significant decrease of 2.97 ($SE = .40$) mean errors over the 12-month period ($p = .000$).

Severity of injury (PTA)

The effect of severity did not reach significance and there was no significant interaction between severity and NART assessment time post injury.

Education

The results showed a highly significant effect of education on NART performance, $F(1, 188) = 23.68, p = .000$, partial $\eta^2 = .11$, $\beta = 1$. Pairwise comparisons showed that overall participants with ≤ 11 years education produced significantly more NART errors ($M = 21.66$, $SE = .82$) than participants with >11 years education ($M = 16.14$, $SE = .79$, $p = .000$). There was a significant interaction between NART assessment time post injury and education as shown in Figure 7.1, suggesting NART recovery was different for each education group $F(2, 376) = 7.21, p = .001$, partial $\eta^2 = .04$, $\beta = .93$.

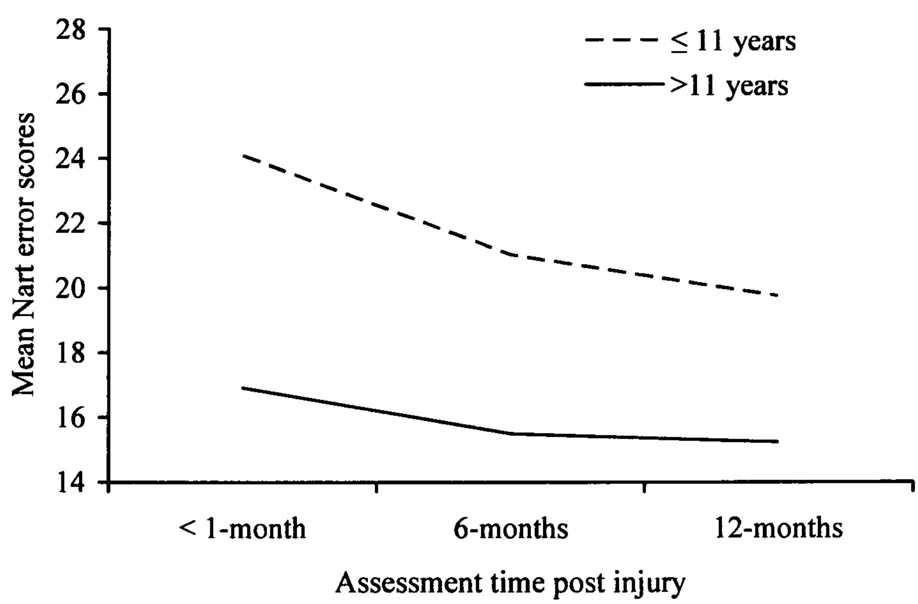

Figure 7.1. Mean NART errors according to years of education for each assessment.

To examine group differences at each time point independent samples t-tests were performed and showed that the group with >11 years of education produced significantly less NART errors than the group with ≤11 years education at the initial assessment, $t(192) = 6.08, p = .000$, the 6-month assessment, $t(192) = 4.81, p = .000$ and at the 12-month assessment, $t(192) = 3.80$,

$p = .000$ Paired samples t-tests showed that the group with ≤ 11 years education produced significantly less mean NART errors by the 6-month assessment, $t(86) = 4.82, p = .000$. They continued to improve and showed significantly less errors at the 12 month assessment compared with the 6-month assessment, $t(86) = 2.16, p = .034$ and an overall significant decrease of 4.30 ($SD = 5.60$) mean errors was observed by the 12 month assessment ($p = .000$). The group with >11 years education also improved by the 6 month assessment, $t(106) = 2.78, p = .007$, producing significantly less mean NART errors compared to the initial assessment but they did not show significant improvement between 6 and 12 months. There was an overall significant decrease of 1.68 ($SD = 5.21$) mean NART errors by the 12 month assessment.

There was a significant three-way interaction between NART assessment time post injury, severity and education, $F(4, 376) = 2.53, p = .04$, partial $\acute{\eta}^2 = .03$, $\beta = .71$. This suggests that the two-way interaction between NART assessment time post injury and education was different for each severity group. Therefore the interaction between NART assessment time post injury and education was examined for each severity group (see Table 7.6a for means and standard deviations and 7.6b for differences between education groups for each severity group at each assessment).

Table 7.6a

Mean (standard deviation) NART errors according to years of education and severity for all three assessments.

Years of Education (*n*)	< 1 month	6-month	12-month
Mild TBI			
≤ 11 years (35)	23.17 (8.49)	20.51 (6.87)	19.86 (8.22)
> 11 years (56)	16.82 (8.03)	14.34 (6.87)	14.93 (7.33)
Moderate TBI			
≤ 11 years (25)	23.32 (9.58)	20.60 (10.55)	18.92 (10.59)
> 11 years (28)	17.18 (5.81)	15.14 (7.00)	14.61 (6.30)
Severe TBI			
≤ 11 years (27)	26.00 (7.25)	22.11 (8.46)	20.41 (8.27)
> 11 years (23)	16.78 (10.14)	18.70 (9.18)	16.74 (10.00)

Table 7.6b

The t values and significance levels for differences in NART errors between education groups for each severity group for all three assessments.

	< 1 month		6-month		12-month	
Severity Group (*df*)	*t*	*p*	*t*	*p*	*t*	*p*
Mild (89)	3.59	.001	4.17	.000	2.98	.004
Moderate (51)	2.78	.008	2.24	.03	1.78[b]	.08
Severe (48)	3.74	.000	1.37	.18	1.42	.16

Note. If Levene's test was significant then the analysis for unequal variances was reported.
[a]df = 38.64 [b]df = 38.17

The group with Mild TBI consisted of 35 participants who completed ≤ 11 years education and 56 participants who completed >11 years education. Independent samples t-tests showed that the group with ≤11 years education produced significantly more mean NART errors than the group with >11 years education at the initial, 6 month and 12 month post injury assessments.

The group with Moderate TBI consisted of 25 participants who attended ≤ 11 years education and 28 participants who attended >11 years education. The results of the independent samples t-tests showed that the group with ≤11 years education produced significantly more mean NART errors than the group with >11 years education at the initial and 6 month assessment but no significant differences were observed at the 12-month assessment. Paired samples t-tests showed a significant decrease in mean NART errors by the 6 month and 12 month assessments for both education groups and no significant change was noted for either group between 6 and 12 months.

The group with Severe TBI consisted of 27 participants who attended ≤ 11 years education and 23 participants who attended >11 years education. Independent samples t-tests showed that the group with ≤11 years education produced significantly more mean NART errors than the group with >11 years education at the initial assessment only. Paired samples t-tests showed that the Severe TBI group with ≤11 years education produced significantly fewer mean NART errors by the 6 and 12 month assessments compared with the initial assessment but no significant change from 6 to 12 months was observed. Participants who had attended >11years education showed no significant change in error scores over time.

Paired samples t-tests showed that both education groups produced significantly less mean NART errors by the 6 month assessment and by the 12 month assessment but no significant change in NART errors was observed between 6 and 12 months. Tables 7.7a shows *t*-values and significance levels for comparisons of NART score at each assessment for each severity group for participants with ≤ 11 years education and Table 7.7b shows these comparisons for participants with >11 years education.

Table 7.7a

The t values and significance levels for comparisons of NART errors according to severity for participants with ≤11 years education (n = 87) for all three assessments.

Severity Group (*df*)	*T*	*P*
Mild		
1 vs 6 (35)	3.05	.004
6 vs 12 (35)	.88	.39
1 vs 12 (35)	4.42	.000
Moderate		
1 vs 6 (25)	2.44	.02
6 vs 12 (25)	1.69	.10
1 vs 12 (25)	3.48	.002
Severe		
1 vs 6 (27)	2.82	.009
6 vs 12 (27)	1.24	.23
1 vs 12 (27)	4.72	.000

Table 7.7b

The t values and significance levels for comparisons of NART errors according to severity for participants with >11 years education (n = 107) for all three assessments.

Severity Group (*df*)	*t*	*P*
Mild		
1 vs 6 (56)	3.59	.001
6 vs 12 (56)	-.83	.412
1 vs 12 (56)	2.63	.01
Moderate		
1 vs 6 (28)	2.70	.01
6 vs 12 (28)	.66	.52
1 vs 12 (28)	3.45	.002
Severe		
1 vs 6 (23)	-1.54	.14
6 vs 12 (23)	1.91	.07
1 vs 12 (23)	.04	.97

7.3.4 Severity & and socio-economic status (SES) analyses

A three-way ANOVA (Severity x SES x Time post injury) with NART score as the dependent variable was performed and SPSS print out for analyses is provided in Appendix B5. A sample of 97 participants was divided into SES and severity groups. The number of participants in each group is provided in Table 7.8.

Table 7.8

Number of participants in each SES group for each severity group.

Severity	Professional	Trades	Semi-skilled	Un-skilled	Total
Mild	7	14	13	8	42
Moderate	13	8	6	1	28
Severe	4	10	9	4	27
Total	24	32	28	13	97

Time post injury

There was a significant main effect of time post injury on NART errors, $F(2, 170) = 6.19$, $p = .003$, partial $\eta^2 = .07$, $\beta = .89$. Mean NART errors was 20.70 ($SE = 0.98$) at the initial assessment, 18.33 ($SE = 0.99$) at the 6-month assessment and 18.60 ($SE = 1.01$) at the 12 month assessment. Pairwise comparisons showed a significant decrease in mean NART errors by the 6 month assessment compared to the initial ($p = .002$) but there was no significant change in NART errors between 6 months and 12 months. An overall significant decrease of 2.10 ($SE = 0.80$) mean NART errors was observed by 12 months ($p = .03$).

Severity of injury (PTA)

The effect of severity failed to achieve an acceptable significance level ($p = .09$) in this analysis and there were no significant two-way or three-way interactions involving severity.

Socio-economic status

There was a significant main effect of SES on NART performance, $F(3, 85) = 7.47, p = .000$, partial $\eta^2 = .20$, $\beta = .98$ (Figure 7.2). The Levene's test was significant and a more stringent statistical significance level was applied, $p < .01$. Pairwise comparisons showed that overall people working in Professional jobs produced significantly less mean NART errors than Trades Persons ($p = .005$), Semi-skilled workers ($p = .000$) and Un-skilled workers. ($p = .013$). To further explore the main effect of SES on NART performance independent samples t-tests were performed at each time point The Professional group showed significantly less NART errors than the Trades group, the Semi-skilled group and the Un-skilled group at all three assessments. The Trade group produced significantly less errors than the Un-skilled group at the 12-month assessment (see Table 7.9 for *t*-values and significance levels). No other significant group differences were observed and there were no significant two-way or three-way interactions.

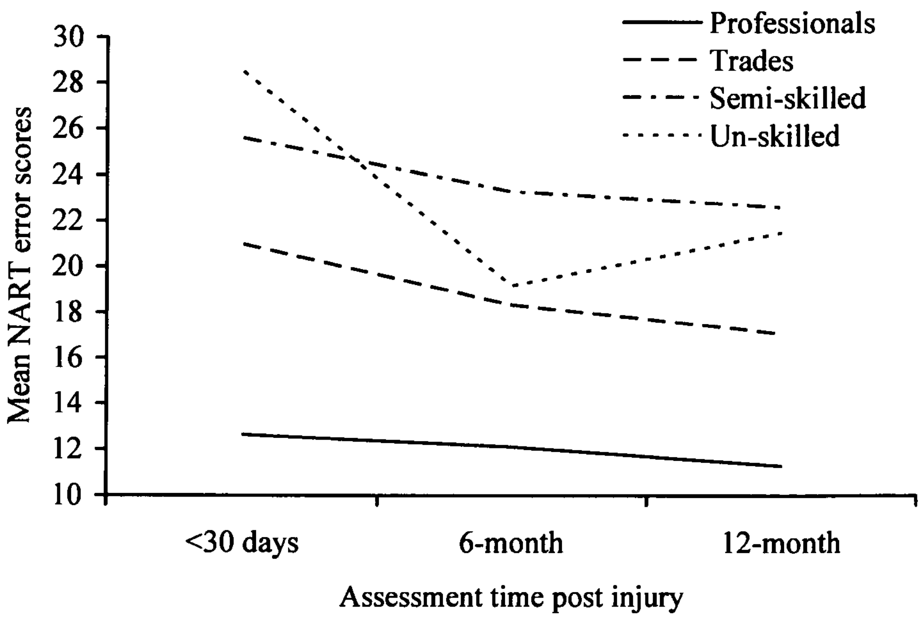

Figure 7.2. Mean NART errors according to SES for each assessment.

Table 7.9

The t-values and significance levels for SES group comparisons of NART errors for each assessment.

	< 1 month		6-month		12-month	
SES Group (*df*)	*t*	*p*	*t*	*p*	*t*	*p*
1 vs 2 (54)	-5.15	.000	-3.28	.002	-3.90	.000
1 vs 3 (50)	-4.70[a]	.000	-4.04[b]	.000	-4.20	.000
1 vs 4 (35)	-4.97[c]	.000	-2.54	.02	-3.94[d]	.001
2 vs 3 (58)	-.76	.45	-1.52[e]	.14	-1.20	.24
2 vs 4 (43)	-1.86	.07	-.40	.70	-2.08	.04
3 vs 4 (39)	-.88	.39	.71	.48	-.90	.37

Note. If Levene's test was significant then the analysis for unequal variances was reported.
[a]df = 41.26. [b]df = 48.05. [c]df = 16.60. [d]df = 15.80. [e]df = 47.14.

7.3.5 Severity & and gender analyses

A three-way ANOVA (Severity x Gender x Time post injury) with NART score as the dependent variable was performed and SPSS print out for all analyses is provided in Appendix B6. The sample was divided into severity and gender groups and the Mild group consisted of 57 males and 34 females. The Moderate group consisted of 27 males and 26 females and the Severe group consisted of 33 males and 17 females.

Time post injury

There was a significant main effect of time post injury on NART errors, $F(2, 376) = 224.25$, $p = .000$, partial $\eta^2 = .11$, $\beta = 1$. Mean NART errors was 20.70 ($SE = 0.98$) at the initial assessment, 18.33 (SE = 0.99) at the 6-month assessment and 18.60 (SE = 1.01) at the 12 month assessment. Pairwise comparisons showed a significant decrease in NART errors by the 6 and 12 month assessments compared to the initial (p = .000) but there was no significant change in NART errors between 6 months and 12 months.

Severity of injury (PTA)

The effect of severity did not reach significance in this analysis and there were no significant two-way or three-way interactions involving severity.

Gender

There was no significant difference between males (M=18.20; SE = .77) and females (M =19.43; SE = .95) NART errors scores and there were no significant two-way or three-way interactions involving gender.

7.4 Discussion

The results supported the hypothesis that the NART can be impaired following traumatic brain injury (TBI). An overall significant improvement in NART performance was observed by 6 months post injury but no significant change in NART score was observed between the 6 and 12month assessments, suggesting NART is sensitive to TBI if administered soon after injury. An overall decrease of 2 NART errors (2 FSIQ points) was observed at the 6-month assessment and an overall decrease of 3 errors (3 FSIQ points) over the 12-month period was observed. Notably, while this was statistically significant the actual change in error score may not be clinically meaningful. Given error is inherent in all cognitive tests and estimates of premorbid functioning, clinicians are required to interpret test results with this in mind. Practice effects could also be posed as possible reasons for the decrease of three NART errors over time. However the change is likely to represent genuine improvement as Crawford (1989b) suggested that practice effects would be of little practical significance, observing a change of less than one error point when examining the test-retest reliability of the NART.

Of greater significance was NART sensitivity in TBI according to years of education, socio-economic status (SES), age and to a lesser degree severity of injury. First, the hypothesis that participants with a severe TBI would make more NART errors than participants with a mild TBI was supported but it was a small effect, and further examination showed that the significant difference occurred only at the 6 month assessment and this was only observed for the severity and age analyses, suggesting NART is fairly robust to severity of injury. No significant difference in NART performance was observed between participants with severe and moderate TBI. Also, while the interaction between NART assessment time and severity did not reach significance visual inspection indicated that the severe TBI group may have been slower to recover than the Mild group and this caused the larger difference in error scores at the 6 month assessment.

Previous findings have been inconsistent but some indicate that NART performance is associated with clinical measures of severity of brain damage (e.g., Morris et al., 2005). The results of the current study showed some support for previous findings suggesting a dose-related effect, with greater injury severity associated with poorer NART performance, although it was a relatively weak finding. The Severe TBI group produced 4 more NART errors (4 FSIQ points) than the Mild TBI group at the 6-month assessment only and this only just reached significance ($p = .04$). Previous studies have utilised correlational designs have yielded inconsistent results and have not fully explored the contribution of severity to NART performance. The findings of the current study suggest other factors such as demographic variables may be more influential on NART sensitivity and recovery following TBI. The present study provides more detailed information on the contribution of severity of injury to NART performance and in doing so adds to the body of literature available.

Contrary to the hypothesis age was found to have a significant effect on NART performance and the results showed that as age increased NART errors decreased. Participants aged 16 – 24 years produced 7 more NART errors than those aged 61 - 80 years at the initial assessment, 5 more errors at the 6-month assessment and 6 more errors at the 12-month assessment.

NART sensitivity and therefore recovery was different across age groups with older TBI patients showing less initial NART sensitivity and therefore, less change in NART performance over time. Although they did exhibit a significant improvement in NART performance over 12 months, the decrease was 2 NART errors only. In contrast younger TBI patients show more sensitivity to NART following TBI, displaying a significant decrease in NART errors by 6 months post-injury.

It is likely that older participants had more familiarity with NART words. The NART words were selected in 1982 and many changes have occurred in the English language making it highly likely that younger people would not have had the same exposure to these words as older people. It is likely that older people had more traditional English education than younger participants resulting in increased familiarity with the NART words.

Previous research examining the influence of age has produced inconsistent results and has often been based on correlational techniques only (Nelson & McKenna, 1975; Crawford, 1988b). Nelson and McKenna found that age did not correlate significantly with the number of NART errors. In contrast Crawford et al. found age did correlate with NART error estimated IQ ($r = -0.18$, $p < .001$) but after partialling out the effects of education and social class the correlation for age was no longer significant, resulting in the authors concluding age did not influence NART performance. In contrast to the previous studies mentioned, utilising correlational designs, the current study provided good evidence for age effects using ANOVA.

The hypothesis that gender would not have an effect on NART performance was supported in this study, which showed no significant differences between males' and females' NART performance, confirming previous research findings (Crawford, 1988b; Wiens et al., 1993).

Education was found to have a highly significant effect on NART performance, as hypothesised, providing further confirmation of previous research, which has found significant correlations between education and NART in normal samples (Crawford et al., 1988b) and TBI samples (Watt & O'Carroll, 1999). The results showed that participants with more years of education produced significantly fewer NART errors than the group with fewer years of education at all three assessments. There was a significant difference of 7 NART errors at the initial assessment, 6 errors at the 6 month assessment and 5 errors at the 12 month assessment.

An interaction between NART assessment time post injury and education was also observed and this suggested that more years of education may provide protection from NART sensitivity to TBI, resulting in minimal change in NART performance over 12 months. Notably, the decrease in NART errors for those with more years of education reached significance but the actual mean change was only 1 NART error. In contrast, those with fewer years of education showed more NART sensitivity to TBI with a significant decline of only 3 NART errors observed by 6 months post-injury.

Additionally, there was a three-way interaction between NART assessment time post injury, education and severity suggesting the interaction was different for each severity group. Those with more years of education show similar initial NART scores regardless of severity. However, recovery varied slightly with mild and moderate injury groups showing a significant decrease of 2 NART errors by 6 months post-injury but no change was observed over time if the injury was severe, suggesting more time may be required for further recovery of the 2 NART errors observed for those with a severe injury. This would be expected given the recovery of other cognitive abilities is differentially affected by severity (Ponsford et al. 1995).

The group with fewer years of education showed similar initial NART performance in people with mild and moderate injuries but in this group, more initial NART errors were produced by the people with severe injury, suggesting NART is sensitive to the severity of injury in this group. Those with fewer years of education produced significantly more initial NART errors than those with longer education at all three assessments and the difference increased with severity, with a difference of 9 NART errors observed between education groups for people with a severe injury. The group with less education showed a significant decrease of 3 errors by the 6 month assessment for all severity groups.

In summary those with fewer years of education show more NART sensitivity to TBI severity than those with more years of education. For participants with fewer years of education, a severe injury resulted in more NART impairment at the initial assessment compared to those with mild or moderate injuries. In contrast, participants with more years of education showed similar initial NART performance regardless of severity of injury but the severe group did not show recovery in NART performance at 12 months as observed for those with mild and moderate injuries who showed a significant decrease of 2 error points over 12 months. Notably for the group with more years of education and a severe injury an increase of 2 errors was observed at 6 months post injury. While this is an unusual result and it is difficult to explain the increase in errors at this time point it did not reach significance. There are very few studies which have examined the relationship between education and NART performance and those that have, have used correlational techniques only. In contrast the present study has utilised ANOVAs due to their ability to provide more detailed information in regards to NART sensitivity and recovery following TBI.

The results of this study have significant implications for the estimation of premorbid IQ in clinical practice suggesting years of education and severity should be considered as they impact on possible NART sensitivity leading to an underestimation of premorbid IQ. This will influence the development of rehabilitation programs and expectations for recovery. It is possible that participants with more years of education have greater cognitive reserve which may protect them from NART underestimates as shown in previous research examining other cognitive abilities (Ander, Vigen, Mack, Clark, & Gatz, 2006; Staff, Murray, Deary & Whalley, 2004). A similar effect of education on NART has been found for patients with dementia. Stebbins et al. (1990) found that the DART (Dutch version of the NART) scores ‘held’ for those with mild dementia only if they had higher education, also supporting the suggestion that education may provide protection from NART sensitivity.

The effect of occupation based socio-economic status (SES) on NART performance was significant, as hypothesised, despite small sample sizes for some SES classes, particularly the Un-skilled Labourers (n = 13). Participants with Professional occupations produced 7 fewer errors than participants working in Trades, 9 fewer than Semi-Skilled workers and 10 fewer errors than Un-skilled labourers. Given the relationship observed between SES and education (Lezak et al., 2004) this finding provides further support for the highly significant effect of education observed in the present study.

There was no significant of severity for the SES analyses, suggesting SES moderates the effect of severity and given the relationship between SES and education it is not surprising to see a similar result for both. Although, given this relationship and the significant interaction observed between NART assessment time, education and severity it is also likely that a similar three-way interaction with SES would have been observed with larger group numbers.

The results showed a differential recovery path of NART performance based on SES with Professionals showing no change in NART errors over time, suggesting little, if any initial

NART sensitivity to TBI. In contrast Un-skilled Labourers produced 6 errors less by the 6-month time revealing NART sensitivity to TBI and reflecting a similar pattern as found for education, that is, both high SES and more years of education provide protection from NART sensitivity to TBI.

Previous research findings have been inconsistent in respect to the effect of SES on NART performance. Crawford et al. (1988b) found that SES did correlate with NART estimated IQ. However, Nelson (1982) concluded that SES did not have a significant independent effect on word reading based on significant differences between mean predicted IQ and mean obtained IQ within each social class. Examining differences in predicted and obtained NART scores for each class fails to assess the contribution of SES to the NART score and the method employed by the current study provides a more valid assessment of the effect of SES on NART performance, supporting Crawford's (1988b) findings.

7.4.1 Limitations

Several potential limitations of the current study may have influenced the results. Firstly, data was obtained from a large population study and a large number (approximately 15) of research assistants administered and scored the NART over several years. Some studies suggest NART possesses high inter-rater reliability and can be reliably administered by inexperienced as well as experienced clinicians (Crawford et al., 1989a; Riley & Simmonds, 2003; Schlosser & Ivison, 1989). However, it has also been observed that raters can differ significantly in the strictness with which they score the NART, despite high correlations between raters (Crawford et al., 1989). Crawford et al. (1989a) also noted that some NART words had low agreement rates. They found that 82% of the words had 90% agreement and 64% had a ≥95% agreement. Words such as *aeon, puerperal, aver, sidereal* and *prelate* yielded the lowest agreement rates. However, it may also be argued that the use of many raters strengthens the robustness of the findings obtained.

The participants in the current study attended all three time points over a one year period suggesting they were motivated to participate in research and that they may not be representative of the general TBI population. Another issue which may have impacted on the effect of severity was that the sample did not include extremely severe TBI participants. The strict inclusion criteria, requiring participants to be assessed within one month of the injury and to be out of PTA may have resulted in some participants in the more severe category being excluded. Examination of the database showed that an additional eight participants with severe TBI (16%) would have been included in the analyses if the initial inclusion criteria had been widened. However, if the inclusion period had been widened then the effect of NART sensitivity for mild and moderate TBI participants may have been lost as recovery would be taking place.

In summary the current study has shown that NART can be impaired following TBI. These findings are relevant to clinical practice, showing NART sensitivity to TBI for people who have less than 12 years education, are from lower SES and are aged between 16 - 24 years. Also of note is that those with severe injuries may take longer than 12 months to show improvement in NART performance. To illustrate these findings, Figure 7.3 provides mean NART errors for those who show the most NART sensitivity, younger people with less than 12 years education, and those who show the least NART sensitivity, older people with more years of education (see Appendix B2 for means).

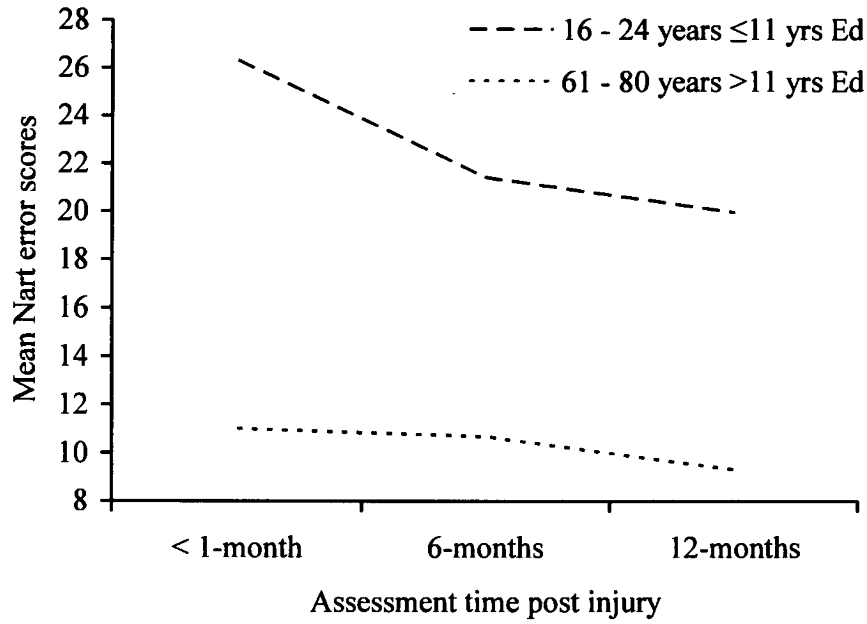

Figure 7.3. Mean NART errors for participants with ≤11 years education aged 16 – 24 years (n = 27) and participants with >11 years education aged between 61 – 80 years (n = 13).

In the cases specified, NART is likely to underestimate premorbid IQ and therefore underestimate cognitive decline, impacting on rehabilitation outcomes and intervention planning. Therefore in neuropsychological practice clinicians will need to consider the influence of factors such as education, age, SES and severity of injury on NART performance when using it as an estimate of premorbid IQ in TBI.

Chapter 8 examines AUSNART performance for participants with TBI, examining the effects of demographic variables and also comparing NART and AUSNART performance in a sub-sample of participants who completed both tests at all three assessments.

Given the findings of Study 1 showing NART sensitivity in TBI, methods for predicting NART performance in TBI will be examined in Chapter 9.

CHAPTER 8

Study 2: Examining the Australian NART (AUSNART) in a TBI Population: Effects of Demographic Variables and Comparison with NART.

The applicability of the British National Adult Reading test (NART) to estimate premorbid IQ in an Australian population has been questioned. It is recommended that NART equations only be used with validity in their country of origin (Hennessey & Mackenzie, 1995). Notwithstanding this, NART norms from the U.K. are generally used for the estimation of premorbid IQ in Australia.

Differences in the educational systems, word pronunciation and accents suggest these British norms may not be applicable for an Australian population (Mathias, Bowden & Barrett-Woodbridge, 2007). Also, several studies have found NART accounts for less variance in FSIQ in Australian samples compared to British samples. Mathias et al. studied a healthy Australian sample ($N = 93$) and found that NART accounted for 44% of the variance in WAIS-R FSIQ.

Hennessey and Mackenzie (1995) found NART accounted for 33% of the variance in WAIS-R FSIQ, and Willshire, et al. (1991) found NART only accounted for 26% of the variance. The variance in FSIQ accounted for by NART appears to be much less in Australian samples, ranging from 26% - 44%, than found in British studies, ranging from 55% - 73% (Crawford et al., 1989a; Crawford et al., 1989d; Crawford et al., 1989c; Crawford et al., 1991; Nelson, 1982). These studies provide further evidence to question the use of NART as an estimate of premorbid IQ in an Australian population.

In response to these concerns the AUSNART was developed by Hennessey and McKenzie (1995) to provide a reliable and valid estimate of premorbid IQ in an Australian population. It was developed on 49 first year University of Tasmania psychology students with an age range of 17 to 40 years (refer to Chapter 4.10, p.69, for more details on the development of the AUSNART). The AUSNART was shown to be a valid and reliable estimate of premorbid IQ revealing a coefficient alpha of .94. Split-half reliability was found to be .94 for the AUSNART showing a slight improvement on previously reported split-half reliabilities for the NART of .90 and .93 (Nelson, 1982; Crawford et al., 1988b). The correlations between the AUSNART and WAIS-R FSIQ, VIQ and PIQ were all significant with AUSNART predicting 55%, 58% and 37% of the variance in WAIS-R FSIQ, VIQ and PIQ. In comparison, while still significant, the NART predicted less IQ variance, yielding 33%, 34% and 18% of the variance for WAIS-R FSIQ, VIQ and PIQ in this sample. Notably as the same participants were used for AUSNART test construction and validation the correlations with IQ would be inflated. However, preliminary results suggest AUSNART may provide a better estimate of premorbid IQ than NART in an Australian population but further large normative studies are required.

There are no published studies examining the AUSNART with large normative samples or with clinical samples. Although a study by Lucas, Carstairs and Shores (2003) examined the validity of a contextual version of the AUSNART (C-AUSNART) in a large normative study comparing seven approaches to estimating premorbid IQ. They placed the AUSNART words in sentences based on research by Beardsall and Hupert (1994) who developed the Cambridge Contextual Reading Test, placing the NART words in context. Beardsall and Hupert suggested people may recognise words better if placed in context and they found an advantage of the contextual presentation over single words for both good and poor readers and it improved the performance of neurological patients. The sample (n = 240) in the Lucas et al. study produced a mean C-AUSNART error score of 25.62 (SD = 11.60) with a range of 4 to 55 and the C-AUSNART accounted for 35% of the variance in WAIS-R FSIQ. Another study by Carstairs, Myors, Shores and Fogarty (2006) examined the influence of language background on tests of cognitive ability including the Contextual version of AUSNART and cautioned against its use with populations of non-English speaking background as it was found to be moderated by language background.

The norms provided for the AUSNART are limited to a small sample of University of Tasmania first year Psychology Students (N = 49) and the results of the study may have been inflated due to the same participants being used for both test construction and validation. Also, it would be beneficial to examine the effects of demographic variables on AUSNART performance or the addition of demographic variables to the AUSNART to estimate premorbid IQ as this was examined in the Hennessey and Mackenzie (1994) study. The utility of the AUSNART with clinical populations, such as TBI, has not been examined and requires research.

8.1 Aims and hypotheses

The study aimed to examine whether the AUSNART is sensitive to TBI and if impairment and recovery are differentially affected by years of education, age and injury severity as found for NART in Study 1 and as observed for other cognitive abilities following TBI (Ponsford et al., 1995). Additionally the study aimed to compare NART and AUSNART performance in an Australian TBI population.

Given there are limited studies examining the AUSNART in the general population and no published research examining the AUSNART in a TBI population, hypotheses were based on findings from Study 1 and the relevant studies reviewed (Hennessey et al., 1995; Lucas et al., 2003; Mathias, Bowden & Barrett-Woodbridge, 2007; Willshire, et al., 1991). The hypotheses were:

- AUSNART performance would be impaired if administered within one month of TBI and show recovery by 6 months, showing a significant decrease in error scores.
- Participants with severe TBI would show significantly fewer AUSNART errors than participants with Mild TBI.
- Mild and Moderate TBI groups would show significantly fewer AUSNART errors by 6 months post injury but the Severe TBI group would not show significantly fewer AUSNART errors until 12 months post injury.
- Younger participants (16 - 24 years) would show significantly more mean AUSNART errors than older participants (61 – 80 years) and they would show significantly fewer NART errors by 6 months compared to the initial AUSNART score. In comparison older participants would not show change in errors over time.

- Participants with ≤ 11 years education would have significantly more mean AUSNART errors than participants who had >11 years education.
- Participants who had ≤ 11 years education would show significantly more AUSNART errors at the initial assessment compared with the 6 month assessment but participants who had >11 years education would not show change in AUSNART performance
- Gender would not significantly influence AUSNART performance.
- NART and AUSNART scores would be significantly correlated but AUSNART would produce significantly higher estimates of FSIQ than NART.

8.2 Method

8.2.1 Participants

A sample of 92 participants with TBI who had completed an AUSNART at 1, 6 and 12 months post injury was obtained from the Neurotrauma Register Research database (NTR). Another sample of 88 participants with TBI, who completed both the NART and AUSNART at all three assessments, was also obtained from NTR to compare performance on the two tests. Please refer to Study 1 for details of the NTR and ethics information.

Demographic data and AUSNART error scores were obtained and participants were selected on the basis they had completed a AUSNART within 1 month of injury and at 6 and 12 months post injury (+ / - 2 months of the 6 and 12 month assessments). Demographic data, AUSNART and NART errors were obtained for a sample of 88 participants who completed both tests at all three assessments as specified above. As in Study 1, length of time in post-traumatic amnesia (PTA) was the measure of severity. The division of the sample into PTA severity groups, age groups and education levels was the same as outlined for Study 1. SES was not analysed as group numbers were too small.

8.2.2 Materials

The Australian National Adult Reading Test (AUSNART; Hennessy & Mackenzie, 1994) is a 64-item phonetically irregular word list presented in order of increasing difficulty (see Appendix C1 for the AUSNART word list). The words cannot be pronounced by common rules of pronunciation such as phonetic decoding. The subject reads the words aloud and error scores are totalled.

The NART (Nelson & Willison, 1991) was also administered to 88 participants who attended all three assessments and is described in Study.

8.2.3 Procedure

Participants were administered the AUSNART word list by trained research assistants (n = 15) at the Neurotrauma Register, as part of a large neuropsychological battery. A sample of 88 participants had completed both the AUSNART and the NART word lists at all three time points as outlined above. Training included instruction in administration and pronunciation of AUSNART and NART words and a tape of the words with the correct pronunciations was given to all research assistants. Standardised administration, as outlined in the NART manual, was adhered to.

Error scores were calculated by Research assistants and all response sheets checked by a Clinical Neuropsychologist. The formulas provided by Hennessey and Mackenzie (1994) were used to convert error scores to estimated FSIQ and these are provided in Appendix A4.

8.2.4 Design and analyses

As for Study 1, a prospective longitudinal, with-in participants design was implemented with AUSNART errors as the repeated measure (dependent variable). Results were analysed using SPSS statistical package (SPSS, 2006). All significance levels were reported at $p = .05$, or better and Bonferroni adjustments made for multiple comparisons.

To examine the influence of and interactions between severity and variables such as education, and age, as observed for NART in Study 1, three-way mixed ANOVAs with AUSNART error score as the dependent variable (repeated measure) were performed. The effect of SES on AUSNART could not be analysed due to small sample size as only 45 participants had information on occupation status. Please refer to Study 1 for details regarding ANOVA analyses.

Correlations of AUSNART and NART errors and estimated FSIQs at each assessment were performed for the sample of 88 participants who completed both tests. To examine NART and AUSNART performance over time (sensitivity) one-way repeated measures ANOVAs were performed for both tests. Given the maximum total score achievable for NART (50 words) and AUSNART (64 words) differed and possible differences in the degree of difficulty of words in each test, paired samples t-tests were performed for the estimated FSIQ scores in preference to comparing the errors scores as this was deemed more clinically relevant.

Data for those who completed an AUSNART is provided in Appendix C1a and data for those who completed both AUSNART and NART is presented in Appendix C1b.

8.3 Results

8.3.1 Descriptive Statistics

The sample (N = 92) was generally representative of a TBI sample in regards to gender, education and age, consisting of 54 males (59%) and 38 females (41 %). Notably there were a marginally higher percentage of females than usual in TBI populations. Mean years of education were 12.7 years and mean age was 39.74. The mean time in PTA was 0.61 of a day (see Table 8.1).

Table 8.1

Descriptive statistics for the total AUSNART sample (N = 92).

	Mean	SD	Median	Range
Age (years)	39.74	17.10	38.44	16 - 77
Years of education	12.71	2.91	12.00	6 – 22
AUSNART1	22.60	11.37	20.00	3 – 48
AUSNART6	21.62	11.10	20.00	3 – 46
AUSNART12	20.27	10.56	18.00	2 – 43
PTA (days)	0.610	1.34	0.08	0 – 7

Motor Vehicle Accidents were the most common cause of traumatic brain injury in this sample, falls and assaults were the second and third most common causes of injury (Table 8.2)

Table 8.2

Frequencies and percentages of cause of injury for the total AUSNART sample (N = 92).

Mechanism	N	Percentage
Motor vehicle accident	31	34
Fall	28	30
Assault	21	23
Sporting accident	6	6.5
Other	6	6.5
Total	92	100

The sample of 88 participants who completed both the NART and the AUSNART at all thee assessments was similar to the other samples studied in the current thesis, consisting of 53 males (60%) and 35 females (40%), The sample (*N* = 88) as a group had a mean age of 38.96 years (*SD* = 17.26) with a range of 16 - 77 years and a mean years of education of 12.82 (*SD* = 2.83), ranging from 6 - 22 years. Mean PTA was 0.55 (*SD* = 1.25) part of a day and ranged from 0 - 7 days and Table 8.3 shows the frequencies and percentages of demographic and clinical variables.

Table 8.3

Frequencies and percentages of demographic/clinical variables for the sample who completed both NART and AUSNART at all each assessment (N = 88).

Demographic Variable	Frequency	Percentage
Severity Group		
Mild	39	44
Moderate	33	38
Severe	16	18
Age Group		
16 – 24 years	27	31
25 – 40 years	21	24
41 – 60 years	31	35
61 – 80 years	9	10
Years of Education		
≤11 years	33	37
>11years	55	63
Cause of Injury		
MVA	30	34
Assault	21	24
Fall	26	29
Sport	6	7
Other	5	6

8.3.2 Severity & age analyses

A three-way ANOVA (Severity x Age x Time Post Injury) with AUSNART as the dependent variable was performed for the 92 participants who completed AUSNART at all three assessments and SPSS printout for analyses are presented in Appendix C3. The total sample was divided into TBI severity groups and age groups as shown in Table 8.4.

Table 8.4

Number of participants in each age group for each severity group.

Severity	16 -24 yrs	25 – 40 yrs	41 – 60 yrs	61 – 80 yrs	Total
Mild	11	10	17	3	41
Moderate	11	6	11	5	33
Severe	4	6	6	2	18
Total	26	22	34	10	92

Time post injury

The results showed a significant main effect of time post injury on AUSNART scores, $F(2, 160) = 7.68, p = .001$, partial $\eta^2 = .08$, $\beta = .95$. Mean AUSNART errors at the initial assessment were 22.37 (SE = 1.40), at the 6 month assessment 21.50 (SE = 1.36) and 19.99 (SE = 1.32) at the 12 month assessment. Pairwise comparisons showed no significant change in AUSNART errors between the initial and 6 month assessments but a significant decrease in AUSNART errors from the 6 month to 12 month assessment was observed (p = .04). Overall, a significant decrease of 2.38 (SE = .68) errors was observed by the 12 month assessment compared with the initial assessment (p = .002).

Severity of injury (PTA)

The results showed no main effect of severity on AUSNART performance and no significant two-way or three-way interactions involving severity of injury.

Age

The effect of age on AUSNART performance showed a trend towards significance ($F(3, 80) = 1.40$, $p = .06$, partial $\acute{\eta}^2 = .09$, $\beta = .62$) and pairwise comparisons showed a trend towards a significant difference between the group aged 16 – 24 years and the group aged 25 – 40 years, with the younger group producing 8.56 ($SE = 3.27$) more errors than the older group ($p = .06$). There were no significant two-way or three-way interactions involving age.

8.3.3 Severity & education analyses

A three-way ANOVA (Severity x Education x Time Post Injury) with AUSNART as the dependent variable was performed and SPSS print out for analyses is provided in Appendix C4. The total sample was divided into TBI severity groups. The sample was divided into TBI severity groups and education groups as shown in Table 8.5.

Table 8.5

Number of participants in each education group for each severity group.

Severity	≤ 11 years	>11 years	Total
Mild	17	24	41
Moderate	10	23	33
Severe	9	9	18
Total	36	56	92

Time post injury

There was a significant effect of time post injury on AUSNART errors, $F(2, 172) = 11.64, p = .000$, partial $\eta^2 = .12$, $\beta = .99$. Mean AUSNART errors were 23.24 ($SE = 1.19$) at the initial assessment, 22.14 ($SE = 1.21$) at the 6 month assessment and 20.45 ($SE = 1.14$) at the 12 month assessment. Pairwise comparisons showed no significant change in AUSNART errors at 6 months compared to the initial assessment but a significant decrease in errors from the 6 month to the 12 month assessment ($p = .01$) and an overall significant decrease in errors over the 12 month period ($p = .000$).

Severity of injury (PTA)

There was no significant effect of severity on AUSNART errors and no significant two-way interactions involving severity. The three-way interaction was not significant.

Education

The results showed a significant main effect of education on AUSNART performance, $F(1, 86) = 12.87, p = .001$, partial η^2 .13, $\beta = .94$. Pairwise comparisons showed that overall participants with ≤ 11 years education produced significantly more AUSNART errors ($M = 26.01$, $SE = 1.72$) than participants with >11 years education ($M = 17.88$, $SE = 1.47$) ($p = .001$). There was a significant interaction between education and AUSNART assessment time post injury, suggesting recovery in AUSNART performance was different for each education group $F(2, 172) = 3.77, p = .03$ partial $\eta^2 = .04$, $\beta = .68$. To further explore the interaction independent samples t-tests were performed and showed that the group with ≤11 years of education produced significantly more AUSNART errors than the group with >11 years education at the initial assessment, $t(90) = 4.57, p = .000$, the 6 month assessment, $t(90) = 3.46, p = .001$ and at the 12 month assessment, $t(90) = 3.36$ $p = .001$ (see Figure 8.1).

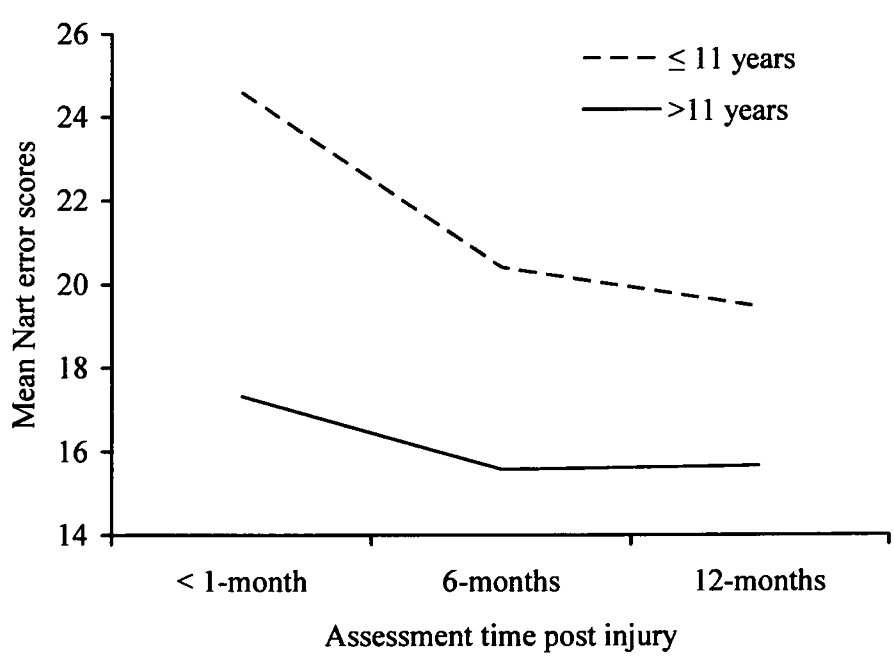

Figure 8. Mean AUSNART errors according to years of education for each assessment.

Paired samples t-tests showed that the group with ≤ 11 years education had significantly less mean AUSNART errors by the 6 month assessment and showed a trend towards significantly fewer errors at the 12 month assessment compared with the 6 month assessment. There was an overall significant decrease in AUSNART errors by the 12 month assessment. The group who had >11 years education showed no significant change in AUSNART performance over time, suggesting they did not experience AUSNART sensitivity to TBI (see Table 8.6 for *t*-values and significance levels). There was no significant three-way interaction involving education

Table 8.6

The *t*-values and significance levels for comparisons of AUSNART errors for each education group for each assessment.

Education Group (*df*)	*t*	*P*
≤11 years (n = 36)		
1 vs 6 (35)	3.08	.004
6 vs 12 (35)	1.97	.06
1 vs 12 (35)	3.71	.001
>11 years (n = 56)		
1 vs 6 (25)	0.12	.91
6 vs 12 (25)	1.78	.08
1 vs 12 (25)	1.82	.07

8.3.4 Severity & gender analyses

A three-way ANOVA (Severity x Gender x Time Post Injury) with AUSNART as the dependent variable was performed and SPSS print out for analyses are provided in Appendix C5. The sample was divided into TBI severity groups and gender groups. The Mild group consisted of 23 males and 18 females. The Moderate group consisted of 18 males and 15 females and the Severe group consisted of 13 males and 5 females.

Time post injury

There was a significant main effect of time post injury on NART errors, with Greenhouse Geisser correction, $F(2, 172) = 9.71$, $p = .000$, partial $\eta^2 = .11$, $\beta = 0.96$. Mean NART errors for the initial assessment were 22.37 ($SE = 1.35$), at the 6 month assessment were 21.17 ($SE = 1.31$) and at the 12 month assessment were 19.69 ($SE = 1.24$). Pairwise comparisons showed no significant change in AUSNART performance between the initial and 6 month assessment but a significant decrease in errors at the 12 month assessment compared to the 6 month assessment ($p = .04$) and an overall significant decrease in errors by the 12 month assessment compared with the initial assessment.

Severity

There was no effect of severity on NART errors and no significant two-way interactions involving severity or three-way interactions.

Gender

There was no significant difference between males (M=20.94, SE = 1.49) and females (M = 21.22, SE = 2.01) NART errors scores.

8.3.5 Comparison of NART and AUSNART performance

To compare NART and AUSNART performance correlations, one-way repeated measures ANOVAs and paired samples t-tests were performed for the sample of 88 participants who completed both tests (details of this sample are provided in sections 8.2 and 8.3 of this Chapter).

Correlations

Pearson correlations (two-tailed) were performed and showed a highly significant relationship ($p < .001$) between NART and AUSNART estimated FSIQ at each time point as shown in Table 8.7.

Table 8.7

Means, standard deviations and correlations (Pearson's) for NART and AUSNART estimated FSIQ.

Assessment	NART	AUSNART	Pearson's Correlation
Initial	103.65 (7.53)	116.77 (9.55)	.88***
6 month	105.05 (7.16)	117.47 (9.26)	.91***
12 month	106.19 (7.01)	118.58 (8.93)	.87***

***$p < .001$

Repeated Measures ANOVAs

The results showed a significant main effect of assessment time post injury on NART performance, $F(2, 174) = 12.17$, $p = .000$, partial $\eta^2 = .12$, $\beta = 0.99$ and on AUSNART performance, $F(1.89, 162.52) = 8.48$, $p = .000$, partial $\eta^2 = .09$, $\beta = 0.96$. Pairwise comparisons showed that NART performance significantly improved by the 6 month assessment ($p = .01$), showing a decrease in NART errors but no significant change in NART performance was observed between 6 and 12 months ($p = .12$). An overall significant decrease in NART errors was observed by 12 months ($p = .000$). In comparison there was no significant decrease in AUSNART errors by 6 months ($p = .29$) but a significant decrease in errors was observed between 6 and 12 months ($p = .03$), resulting in an overall significant decrease in AUSNART errors by 12 months ($p = .000$) as observed for the NART.

Results of t-tests

Paired samples t-tests showed that AUSNART estimated FSIQ was significantly higher than NART estimated FSIQ at all three assessments as shown in Table 8.8.

Table 8.8

Means (standard deviations) and t-values for comparisons of NART and AUSNART estimated FSIQs for each assessment.

Assessment	NART	AUSNART	t-value
Initial	103.65 (7.53)	116.77 (9.55)	-26.69***
6 month	105.05 (7.16)	117.47 (9.26)	-28.72***
12 month	106.19 (7.01)	118.58 (8.93)	-25.75***

***$p < .001$

8.4 Discussion

The results did not support the hypothesis that AUSNART would show sensitivity to TBI with a decrease in errors observed at 6 months post injury as observed for the NART in Study 1. Significant improvement was observed for AUSNART between 6 and 12 months and an overall improvement of 2 errors was observed over the 12 month period. The results suggest that overall AUSNART shows similar NART sensitivity to TBI but takes longer to show improvement as a significant decrease in AUSNART errors was not observed until 12 months post injury.

Also, contrary to the hypothesis AUSNART performance was not influenced by injury severity and therefore, the finding of Study 1 showing NART sensitivity to injury severity was not supported in the current study. The results suggest that AUSNART is not sensitive to the severity of TBI which is in contrast to studies with the NART and the Wechsler Test of Adult Reading (WTAR: Wechsler,) which show dose-related effects, with greater TBI severity associated with poorer performance (Mathias, Bowden, Bigler et al., 2007; Morris et al., 2005).

Notably, NART did not consistently show severity effects in Study 1 as demographic variables such as education, SES and age appeared to be of greater influence and the effect of injury severity on NART performance did not reach significance for the analyses with education and SES. When severity was analysed with age it only just reached significance ($p = .04$) for NART. Together these results suggest that demographic variables, particularly education, SES and age have a more significant effect on AUSNART and NART performance than severity of injury.

The hypotheses in regards to education were fully supported as AUSNART performance and recovery over time were observed to be differentially affected by years of education. People with higher education (>11 years) produced significantly fewer AUSNART errors at the initial, 6 and 12 month assessments (7, 6 and 5 AUSNART errors respectively).

Interestingly, the group with more years of education showed no change in AUSNART performance over time but the group with less years of education showed a significant decrease of 2 errors by 6 months and another decrease of 2 errors by 12 months. This suggests that participants with more education did not show AUSNART sensitivity to TBI but participants with less education did, supporting the findings of Study 1, although to an even greater degree. These findings also provide further support for the suggestion that more highly educated people may have cognitive reserve which protects them from NART and AUSNART sensitivity to TBI as has been noted in research examining other cognitive abilities (Ander, Vigen, Mack, Clark, & Gatz, 2006; Staff, Murray, Deary & Whalley, 2004) and this will be discussed in the general discussion.

This finding has significant implications for the estimation of premorbid IQ in clinical practice suggesting years of education will significantly impact on the expected degree of AUSNART sensitivity to TBI, and subsequent recovery of AUSNART performance, thereby influencing expectations for recovery and rehabilitation program development.

The hypothesis that age would significantly affect AUSNART performance was not fully supported as only a trend towards a significant effect was observed. The young group aged 16 – 24 years produced the most errors and showed a trend towards significantly more errors that the group aged 25- 40 years. Notably, there were small group numbers and limited power to detect an effect (β = .62). However it is also possible that the AUSNART words, developed in 1994 with an Australian population, compared to the NART, developed with a British population in 1982, are more familiar to younger Australians with TBI than the NART words, as the youngest participants in the study were born in 1989. In comparison the NART words were more familiar to older participants as language, particularly vocabulary, has changed significantly since the NART was developed resulting in many younger participants having limited exposure to the NART words but more familiarity with the AUSNART words. Also, the AUSNART words were developed with younger participants aged 17 – 40 years and hence are more appropriate and familiar for this population.

The hypothesis that gender would not have an effect on NART performance was supported in this study, showing no differences between males and females NART performance, as observed in Study 1 and in a previous study examining the influence of demographic variables on NART performance (Crawford, 1988b; Wiens et al., 1993).

The results supported the hypothesis that NART errors and estimated FSIQ would be significantly correlated with AUSNART errors and estimated FSIQ, with correlations, ranging from .80 to .92 observed. The hypothesis that AUSNART would produce higher estimates of FSIQ than NART was also supported with AUSNART producing estimates of FSIQ 13 points higher than the NART estimated FSIQ (FSIQ = 119 & 106 respectively).

However, this finding should be treated with caution as there are several possible explanations for the AUSNART estimated IQ scores to be in the high average range; first, given this sample had a high percentage of participants who attended school for >11 years (63%) it would be expected that this sample would have a higher mean FSIQ than the general population FSIQ mean of 100.

Secondly, the AUSNART equations used in the current study were developed by Hennessey & and McKenzie (1994) using a very small above average IQ sample. The same sample was used for both test construction and validation and to establish the final version's correlation with IQ and this would have inflated the AUSNART estimates of FSIQ.

8.4.1 Limitations

Several limitations may have affected the results. It may be argued that improvement in AUSNART performance could be explained by practice effects and given the very limited research conducted on the AUSNART this is difficult to discount. As Practice effects have been found to be of limited clinical significance for the AUSNART possesses some similar characteristics to the NART it is likely practice effects will be of limited clinical significance for AUSNART also. Also, the analyses in the current study, showing AUSNART sensitivity according to years of education supports the premise that practice effects are of little significance for the AUSNART. The higher education group showed no AUSNART sensitivity to TBI resulting in no significant change in performance over time and hence no evidence of practice effects. In contrast those with fewer years of education showed a decrease of four error points over time suggestive of a true effect and not practice effects.

The sample may not be truly representative of a TBI sample as the sample consisted of larger percentage of participants with >11 years education (56%) than generally observed in TBI samples and a high percentage of participants who were of at least average intellectual functioning premorbidly with 97% of the sample scoring ≤ 43 AUSNART errors (FSIQ = 100) at the initial and 6 month assessments and 100% scored ≤ 43 errors (FSIQ = 100) at the 12 month assessment. The mean AUSNART estimated FSIQ at the 12 month assessment was 118.58 (range of 98– 134) and in the high average range.

The participants in the current study attended all three time points over a one year period suggesting they were motivated to participate in research and may not be truly representative of a TBI population.

The current study provides detailed information regarding AUSNART performance in an Australian TBI sample, detailing the influence of demographic variables on AUSNART sensitivity to TBI and recovery. AUSNART appears to have similar characteristics to the

NART, it is affected by TBI with a decrease in errors observed over time and education significantly affects performance, but further large normative studies and studies examining TBI are required before it can be used with confidence. Further studies of the AUSNART in normative and clinical samples will aide interpretation of the results of the current study.

In summary, the current study has shown that AUSNART can be sensitive to TBI, but only for participants who have had less than 12 years education. In such cases AUSNART is likely to underestimate premorbid IQ by approximately 4 error points and therefore underestimate cognitive decline which will impact on the development of interventions and rehabilitation outcomes.

CHAPTER 9

Study 3: Predicting NART performance in Traumatic Brain Injury

The main aim of this thesis has been to examine if NART and AUSNART estimates of premorbid IQ are sensitive to TBI. The findings from Study 1 indicated that NART can be sensitive to TBI, suggesting it would be clinically useful to examine methods of predicting NART score from the initial 'impaired' NART score and demographic variables. As further research is required before AUSNART can be used with confidence in an Australian TBI population prediction methods were developed only for the NART in the current study as it continues to be the most commonly used estimate of premorbid IQ in neuropsychological practice.

As reviewed in Chapter 3, demographic equations using age, gender, occupation, race and education have been developed to predict premorbid intelligence and have been found to account for between 36% and 53% of the variance in FSIQ (Barona et al., 1984; Crawford et al., 1989d; Wilson et al., 1978). While these have the advantage of being totally objective and not requiring the judgement of the clinician, questions regarding the predictive power of this procedure, regression towards the population mean, restricted IQ range and large standard error of measurement have limited the utility of the purely demographic based regression equations (Langeluddecke & Lucas, 2004; Vanderploeg, Schinka & Axelrod, 1996).

Combined demographic and current measure equations have also been developed to estimate premorbid IQ. These equations have combined demographic variables, such as education and age, with 'hold tests' such as WAIS Vocabulary or reading estimates such as the NART. Research findings have been inconsistent with some studies reporting increased variance in FSIQ accounted for by combined equations compared to purely demographic equations or 'hold test' equations but others studies have not found this to be the case. Crawford et al. (1989d) found that equations incorporating NART and demographic variables (age, gender and occupation/social class) accounted for 73%, 78% and 39% of the variance for WAIS FSIQ, VIQ and PIQ respectively but education did not significantly increase the amount of variance. Berry et al. (1994) also found the addition of demographic variables to NART-R increased the amount of variance explained in FSIQ for a US sample but in contrast they found education and age were the important demographic predictors. Willshire et al. (1991) found the inclusion of education with the NART provided a substantially higher estimate of premorbid cognitive functioning than NART alone in an Australian sample. The combined equations of NART error score and education accounted for 46% of the variance in WAIS-R IQ in the total sample compared to only 26% of the variance in WAIS-R IQ accounted for by NART error score alone. Interestingly, the equation accounted for a much higher percentage of the FSIQ variance (67%) for the older age group (>55years).

Conversely, Blair and Spreen (1989) studied the utility of combining demographic variables to their adaptation of the NART, the New Adult Reading Test (NART-R) for North Americans and found that adding demographic variables did not significantly increase the percentage of explained variance in predicted FSIQ. Bright, Jaldow, and Koleman (2002) also found no advantage of the combined demographic-NART equation in a UK sample comparing controls and a mixed neurological sample. The results suggest that the predictive power of the equations may vary depending on variance specific to the sample studied and therefore suggest that prediction equations should be developed with a large TBI sample if they are to be used to assist clinicians estimate premorbid IQ for patients with TBI.

As reviewed in Chapter 5, prediction equations have been developed to account for NART sensitivity to dementia severity. Taylor et al. (1996) examined the validity and stability of the American version of the NART (AMNART) to estimate premorbid VIQ for patients with Dementia of the Alzheimer's Type (DAT). The authors produced a correction factor based on performance on the Mini Mental Status Examination (MMSE), an independent measure of global cognitive impairment, to account for the negative effects of increasing dementia severity on AMNART performance. The correction factor is clinically useful in the assessment of premorbid IQ for persons with dementia.

To address the concerns of NART sensitivity to brain injury Crawford et al. (1990a) developed an equation to predict a person's NART score using demographic variables, including education, class, age and gender and comparing this with the obtained NART score (see review in Chapter 4.7). The equation has not been used extensively in research, although findings from a study by Skilbeck, et al. (2005) supported its use as a valid estimate of NART performance in a UK neurological sample (n = 175). The authors suggested that the risk of 'impaired' NART performance is greater in patients with diffuse brain damage and lower premorbid VIQs. However, Riley and Simmonds (2003) suggested that it is not entirely clear how accurate a demographically based equation, developed on a neurologically intact sample, is going to be in TBI. The TBI population has a higher rate of premorbid difficulties, including poor academic performance and substance abuse, and often have a lower premorbid IQ than their demographically-equivalent counterparts and Riley and Simmonds suggest the equation may overestimate premorbid IQ in the TBI sample rather than reflect impairment of NART performance. Therefore an equation developed with a TBI sample will provide a more valid and accurate estimate of NART sensitivity, or 'impairment', expected for individuals who have experienced TBI. Further to this, a prediction equation developed with TBI participants from a large population study will provide clinicians with a useful and practical way of calculating a more accurate NART score leading to a more accurate estimate of premorbid IQ for the individual with TBI.

9.1 Aims & hypothesis

The aim of the current study was to develop a prediction method to account for NART sensitivity to TBI, to assist clinicians in making a more accurate estimation of premorbid IQ in individuals with TBI.

It was hypothesised that a prediction equation using the combined NART initial error score and demographic variables, education and age, as predictors would provide a better prediction of the recovered 12 month NART error score than equations utilising either demographic or NART scores alone.

9.2 Method

9.2.1 Participants

The sample consisted of the 194 TBI participants from Study 1 who had completed a NART within 1 month of injury and at 6 and 12 months post injury (+ / - 2 months of the 6 and 12 month assessments). The measure of severity was PTA and NART error scores and demographic data were obtained. Details are provided in Study 1 (Chapter 7.2, p. 97).

9.2.2. Materials

The National Adult Reading Test (NART; Nelson & Willison, 1991) was administered to participants at all three time-points. The NART is a 50-item phonetically irregular word list presented in order of increasing difficulty (Appendix A1). The words cannot be pronounced by common rules of pronunciation such as phonetic decoding. The subject reads the words aloud and error scores were calculated.

9.2.3 Procedure

Participants were administered the NART word list in the same manner as described in Study 1.

9.2.4 Design and Analyses

Results were analysed using SPSS statistical package (SPSS Graduate Pack 15.0 for Windows, 2006). Correlations between demographic variables (years of education, age, PTA, gender) and both the initial 'impaired' NART error score and the 12 month 'recovered' NART error score were performed. Multiple Regression analyses were performed to examine the amount of variance accounted for in NART by demographic variables alone and by combined methods using demographic variables and NART. Predictors (age, education, gender, severity and NART initial error score) were entered in a hierarchical manner using the forced entry procedure in SPSS. SES was not included in the multiple regression equation as only half the sample had occupation information and it was preferable to use the total sample of 194 participants for multiple regression. All significance levels were reported at p =.05, or better, and Bonferroni adjustments made for multiple comparisons. Data for Study 3 is provided in Appendix D1.

9.3 Results

9.3.1 Descriptive Statistics

As the sample used for the multiple regression analyses is the sample used in Study 1 details will not be presented again but are provided in Chapter 7 (p. 98)

9.3.2 Correlations

The results of the correlational analyses are displayed in Table 9.1. The correlations for NART and estimated FSIQs with other variables are identical as would be expected given that estimated IQs are linear transformations of NART scores. The results showed that education was highly correlated with the initial NART and estimated FSIQ accounting for 24% of the variance in each. Education was also significantly correlated with the 12 month NART errors and the estimated FSIQ but to a lesser degree than for the initial scores, accounting for 17% of the variance in each. Age was significantly correlated with NART initial error and estimated FSIQ accounting for 8% of the variance in each. Similarly, age was correlated with the 12 month NART errors and estimated FSIQ but to a lesser degree, accounting for 6% of the variance in each. NART initial and 12 month errors were highly significantly correlated with the initial and 12 month estimated FSIQs (see Appendix D2).

Table 9.1

Correlations of NART scores and estimated FSIQs with education and age for the total sample (N = 194).

	Yrs Ed	Age	NART 1	NART 2	FSIQ 1	FSIQ 12
NART 1	-.49***	-.29***	1	.75***	-.1***	-.75***
NART 12	-.42***	-.24**	.75***	1	-.75***	-1***
FSIQ 1	.49***	.29***	-1***	-.75***	1	.75***
FSIQ 12	.42***	-.24**	-.75	-1***	.75***	1

Note. 1 = initial assessment conducted ≤1 month of injury; 12 = 12 month post injury assessment
$^{**}p < .01$ $^{***}p < .001$.

9.3.3 Multiple Regression

Multiple Regression analyses were performed to examine the amount of variance in NART errors accounted for by a demographic equation and SPSS print out is presented in Appendix D3. Exploratory analysis with years of education, age, severity (PTA) and gender entered as predictors and NART initial error as the outcome variable showed that education and age significantly contributed to the model but gender ($t = 1.45$, $p = .148$) and severity ($t = .31$, $p = .76$) did not.

A multiple regression equation with education and age entered as predictors and NART initial error as the outcome variable showed that education accounted for 24% of the variance in NART initial error and when age was included the model accounted for 33% of the variance in

NART initial errors. Table 9.2 shows the individual contributions of education and age to the model.

Table 9.2

Multiple Regression for predicting NART initial errors from demographic variables.

		B	*SE B*	*B*	*T*	*Sig*
Model 1						
	Constant	39.67	2.55		15.53	***
	Education	-1.59	.203	-.49	-7.84	***
Model 2						
	Constant	45.63	2.68		17.06	***
	Education	-1.62	.19	-.50	-8.44	***
	Age	-.147	.03	-.30	-5.08	***

Note. R^2 = .24 for Model 1 (p < .001); Δ R^2 for Model 2 = .09 (p < .001)
***p < .001.

The model accounted for a lower percentage of the variance in 12 month NART errors than for the initial NART errors with education accounting for 17% of the variance in each at 12 months. When age was included the model accounted for 23% of the variance. The individual contribution of education and age to the model are provided in Table 9.3.

Table 9.3

Multiple Regression for predicting 12 month NART errors from demographic variables.

		B	*SE B*	*Β*	*T*	*Sig*
Model 1						
	Constant	33.07	2.55		12.95	***
	Education	-1.29	.203	-.42	-6.34	***
Model 2						
	Constant	37.95	2.75		13.79	***
	Education	-1.31	.20	-.42	-6.67	***
	Age	-.118	.03	-.25	-3.96	***

Note. $R^2 = .17$ ($p < .001$); ΔR^2 for Model 2 = .06 ($p < .001$)
***$p < .001$

To produce a combined demographic and NART equation to predict the improved 12 month NART error score, NART errors, years of education and age were entered as predictors (Table 9.4).

Table 9.4

Multiple Regression for predicting 12 month NART errors from the initial NART error scores and demographic variables.

	B	*SE B*	*β*	*T*	*Sig*
Model 1					
Constant	1.86	.91		2.04	*
NART Initial Error	.765	.04	.80	18.46	***
Model 2					
Constant	3.26	2.54		1.29	ns
NART Initial Error	.75	.05	.79	15.75	***
Education	-.09	.15	-.03	-.59	ns
Model 3					
Constant	3.84	2.99		1.29	ns
NART Initial Error	.75	.05	.78	14.62	***
Education	-.10	.16	.03	-.65	ns
Age	-.01	.02	.02	-.37	ns

Note. $R^2 = .64$ for Model 1 ($p = .000$); ΔR^2 for Model 2 = .001 (ns); ΔR^2 for Model 3 = .00 (ns)
*$p < .05$. *** $p < .001$.

The results showed that NART initial errors predicted 64% of the variance in the 12 month improved NART error. The addition of education and age did not significantly improve the prediction equation. Therefore, the prediction equation adopted was as follows:

Improved 12 month NART Error = 1.861 + (initial NART error score x .765)

Given the equation accounted for a significant amount of variance in the improved 12 month NART error score it was surprising that education and age did not significantly add to the predictive ability given the significant effect of education and age on NART performance observed in Study 1. Therefore, Multiple Regression was performed for each education group to examine if the equation would account for more of the variance in the 12 month NART. However, the equations did not increase the predictive ability of the model, accounting for 63% of the variance in NART 12 month errors for the group with less years of education and 61% of the variance for those with more education (see Appendix D4 and D5). Therefore the equation using the NART initial error score only as the predictor was the preferred prediction method. A table providing the conversion of obtained NART errors to predicted NART error scores using this equation is provided in Appendix A4.

9.3.4 Accuracy rates for the NART prediction equation

To examine the accuracy of the prediction equation a NART predicted – obtained score was calculated for each participant and SPSS print out is provided in Appendix D6. Overestimating (negative score) suggested the person was getting better than the actual obtained NART score shows and indicates there has been a greater decrease in errors than actually occurred, hence more improvement than actually observed. Conversely, underestimating (positive score) suggests the person is getting worse than the actual obtained NART error score shows and therefore indicates there is an increase in errors which is greater than what is actually observed, hence less improvement than is actually observed. A score of 0 indicated the prediction equation provided an accurate estimate of the 12 month obtained NART score. To allow for normal variations which may occur (Crawford, 1989a) scores between -1 and +1 were classified as accurate predictions. Table 9.5 shows percentages of predicted - obtained difference scores observed for the sample.

Table 9.5

Percentages and frequencies (n) of predicted - obtained difference scores.

Predicted – Obtained Difference Scores									
≥ -8	-7→-6	-5→ -4	-3→ -2	**-1 →+1**	2 → 3	4 → 5	6 → 7	≥8+	Range
6% (12)	7% (14)	11% (21)	19% (38)	**17% (34)**	12% (23)	15% (29)	7% (11)	6% (12)	-14 → 15

The results showed that the prediction equation accurately predicted the 12 month NART score in 17% of cases. Notably, the equation predicted the improved 12 month NART within + - 3 errors for 48% of the cases.

To examine the accuracy of the equation according to IQ the sample was split into two groups based on the 12 month estimated FSIQ: Group 1 consisted of participants with a FSIQ of ≤ 100 and Group 2 consisted of participants with a FSIQ >100. Table 9.6 shows accuracy rates for the initial NART error prediction equation for each IQ group and SPSS print out is provided in Appendix D7.

Table 9.6

Percentages and frequencies (n) of predicted - obtained difference scores according to FSIQ for the NART initial error prediction equation.

	Predicted – Obtained Difference Scores								
	≥ -8	-7→-6	-5→ -4	-3→ -2	-1→+1	2 → 3	4 → 5	6 → 7	≥8+
FSIQ ≤ 100 (44)	18% (8)	16% (7)	18% (8)	30% (13)	**11% (5)**	5% (2)	2% (1)	-	-
>100 (150)	3% (4)	5% (7)	9% (13)	17% (25)	**19% (29)**	14% (21)	19% (28)	7% (11)	8% (12)

The range of predicted – obtained scores for the group with a FSIQ ≤ 100 was from -14→ 4 and from -11 → 15 for the group with a FSIQ > 100. The equation accurately predicted the 12 month NART in 19% of cases for the high IQ group compared with 11% of cases in the low IQ group. However, the equation predicted the 12 month NART within +- 3 error points for 46% of the low IQ group and 50% of the high IQ group.

9.3.5 Mean change score according to education and age

An alternate way to calculate estimates of the 12 month NART was to examine mean change scores between the initial and 12 month NART assessments for each age group within each education level. Means and standard error for the initial impaired NART errors score, the recovered 12 month NART and the mean improvement in NART errors over time are presented in Table 9.7 and SPSS print out is provided in Appendix D8.

Table 9.7

Means (standard error) for the initial impaired NART, 12 month recovered NART and mean NART improvement over 12 months for each education level and each age group.

Years of Education (n) Age (n)	Initial Impaired NART	12 month Recovered NART	Improvement
≤ 11 Years (87)			
16 – 24 years (27)	26.30 (1.41)	19.96 (1.60)	6.33 (1.04)***
25 – 40 years (20)	24.20 (1.76)	20.00 (2.25)	4.20 (1.51)*
41 – 60 years (25)	23.92 (1.87)	20.56 (1.61)	3.36 (1.09)**
61 – 80 years (15)	20.27 (2.31)	17.73 (2.60)	2.53 (0.96)*
> 11 years (107)			
16 – 24 years (39)	21.13 (1.33)	19.56 (1.28)	1.56 (.883)
25 – 40 years (26)	17.15 (1.09)	14.62 (0.95)	2.54 (1.16)*
41 – 60 years (29)	13.66 (1.25)	12.62 (1.35)	1.03 (0.88)
61 – 80 years (13)	11.00 (1.96)	9.31 (1.64)	1.69 (1.10)

*$p < .05$. **$p < .01$. ***$p < .001$.

Overall, individuals with fewer years of education experience more NART sensitivity following TBI than individuals with more education and people in the lower education group aged between 16 – 24 years of age display the most sensitivity, showing a mean decrease of 6.33 errors by 12 months post-injury. While NART scores have been reported in the current thesis Table 9.8 shows the corresponding estimated FSIQ scores.

Table 9.8

Means (standard errors) for the initial estimated FSIQ, 12 month estimated FSIQ and mean IQ improvement over 12 months for each education level and each age group.

Years of Education(n)	Initial Estimated FSIQ	12 month Estimated FSIQ	Improvement
≤ 11 Years (87)			
16 – 24 years (27)	96.26 (1.49)	102.81 (1.67)	6.56 (1.07)***
25 – 40 years (20)	98.50 (1.84)	102.80 (2.33)	4.30 (1.58)*
41 – 60 years (25)	98.76 (1.95)	102.08 (1.69)	3.32 (1.15)**
61 – 80 years (15)	102.47 (2.41)	105.07 (2.68)	2.60 (0.97)*
> 11 years (107)			
16 – 24 years (39)	101.67 (1.38)	103.28 (1.32)	1.62 (0.90)
25 – 40 years (26)	105.77 (1.12)	108.38 (0.95)	2.62 (1.18)*
41 – 60 years (29)	109.24 (1.26)	110.31 (1.38)	1.07 (0.89)
61 – 80 years (13)	112.00 (1.96)	113.69 (1.64)	1.69 (1.10)

*$p < .05$ **$p < .01$ ***$p < .001$

9.3.6 Accuracy rates for the mean change prediction method

To examine accuracy rates of the mean change method of predicting NART performance in TBI, percentages of predicted – obtained difference scores were calculated and are presented in Table 9.9 and SPSS print out is provided in Appendix D8.

Table 9.9

Percentages and frequencies (n) of predicted – obtained difference scores for each education level and age group according to the mean change prediction method.

Education Age	≥ -8	-7→-6	-5→-4	-3→-2	**-1→1**	2 → 3	4 → 5	6 → 7	≥8+
				Predicted – Obtained Difference Scores					
≤11 years (n = 87)									
16 – 24yrs (27)	7% (2)	4% (11)	19% (5)	7% (2)	**19% (5)**	15% (4)	7% (2)	11% (3)	11% (3)
25 – 40yrs (20)	10% (2)	10% (2)	5% (1)	25% (5)	**15% (3)**	10% (2)	15% (3)	-	10% (2)
41 – 60yrs (25)	4% (1)	8% (2)	4% (1)	32% (8)	**20% (5)**	12% (3)	4% (1)	4% (1)	12% (3)
61 – 80yrs (15)	-	13% (2)	7% (1)	27% (4)	**20% (3)**	13% (2)	20% (3)	-	-
>11 years (n = 107)									
16 – 24yrs (39)	13% (5)	8% (3)	5% (2)	8% (3)	**23% (9)**	23% (9)	10% (4)	8% (3)	2% (1)
25 – 40yrs (26)	8% (2)	12% (3)	15% (4)	15% (4)	**19% (5)**	8% (2)	12% (3)	-	12% (3)
41 – 60yrs (29)	3% (1)	7% (2)	14% (4)	17% (5)	**24% (7)**	14% (4)	3% (1)	14% (4)	3% (1)
61 – 80yrs (13)	-	8% (1)	15% (2)	23% (3)	-	46% (6)	8% (1)	-	-

Accurate predictions (-1 → +1 predicted – obtained difference scores) of the 12 month NART ranged from 15% accuracy for the group aged 25 – 40 years with fewer years of education to 24% accuracy for the group aged 41 – 60 years with more years of education. For participants with >11 years education the 12 month NART was predicted within +- 3 error points for 54% of cases aged 16 – 24 years; 42% of cases aged 25 40 years and 55% of participants aged 41 – 60 years. For those aged between 61 – 80 years the mean change prediction method did not accurately (+ -1error points) predict the 12 month NART for any cases but predicted did predict it within +- 3 errors points for 69% of this group.

For participants with ≤11 years education the 12 month NART was predicted within +- 3 error points for 41% of cases aged 16 – 24 years, 50% of cases aged 25 – 40 years, 64% of cases aged 41 – 60 years and 60% of cases aged 61 – 80 years.

9.4 Discussion

The current study examined two methods of predicting NART score, given its observed sensitivity soon after TBI: a regression equation using the initial NART error score and a mean change score according to education and age. As noted in research reviewed prediction equations can be clinically useful in the estimation of premorbid IQ, particularly in circumstances where NART performance may be impaired, such as dementia and TBI.

The hypothesis that a combined initial NART score and demographic equation would provide a better prediction of the recovered NART error score than equations based on either demographic variables or NART scores alone was not supported. Contrary to expectations, the addition of age and education to the initial NART error prediction equation did not significantly increase the amount of variance accounted for in the 12 month NART, suggesting the variance accounted for by education and age were explained in the initial NART error score. To further examine this prediction equations were developed for each education group but this did not significantly improve the equation accounting for 63% of the variance in the 12 month NART for those with ≤11 years of education and 61% for those with more than >11 years.

The preferred equation using the initial NART error score accounted for 64% of the variance which is relatively high compared with other prediction equations. For example, Crawford et al. (1990a) developed an equation using demographic variables to predict NART score in cases of TBI or dementia when the obtained NART was thought to be impaired and the equation only accounted for 49% of the variance in NART error score. In the current study the amount of variance accounted for by education and age was found to be less for the 12 month NART than for the initial NART. Prediction equations developed for each education group did not significantly improve the equation accounting for 63% of the variance in the 12 month NART for those with ≤11 years and 61% for those with > 11 years.

Education was found to account for 24% of the variance of the initial NART error score and when age was also entered an additional 9% of the variance was accounted for, resulting in a total of 33% of the variance accounted for in initial NART error. For the 12 month NART education accounted for 17% of the variance and age accounted for 6%, together accounting for 23% of the variance in NART 12 month error score. This suggests that education, and to a lesser degree age have a significant influence on NART sensitivity to TBI. Notably the amount of variance predicted from demographic variables (education and age) lessens over time and by 12 months the equation accounts for approximately two-thirds of the variance observed soon after injury. However, if NART cannot be administered due to dyslexia or dysphasia then this method can assist clinicians predict premorbid IQ.

In contrast, Crawford et al. (1990a) found that demographic variables (education, age, class and gender) accounted for 49% of the variance in NART score but this was developed on a sample of healthy participants. The NART score obtained at 12 months post injury in the current study was presumably at a time when the participant had recovered substantially from the TBI and should therefore represent a score comparable to the participants in the Crawford et al. study. Interesting, at 12 months post injury the variance accounted for by demographic variables is not only less than that accounted for soon after the injury but much less than what was found in the Crawford et al. study: 23% versus 49%. This suggests the amount of variance accounted for by education and age is substantially different in TBI populations compared to neurologically intact samples, providing support for Riley and Simmonds (2003) suggestion that it may not be accurate to use a demographically based equation developed on a neurologically intact sample (Crawford et al., 1990a) for individuals with TBI. While Crawford also included SES in the demographic equation and due to a small sample with SES information this was not possible in the current study, it is not likely to have significantly increased the amount of variance accounted for to similar to that found in the Crawford et al. study as education and SES are likely to co-vary.

However, another study by Crawford et al. (1989d) found no additive effect of education on the proportion of IQ variance predicted with a UK sample without neurological impairment and concluded this was due to the amount of covariance between NART and education. The

Crawford et al. study found that the NART-demographic combined equation (NART + age + gender + SES) accounted for 73% of the obtained WAIS FSIQ, NART alone accounted for 66% of the variance and demographic variables alone accounted for 50% of the variance, hence an improvement in the combined equation not observed in the current study. Notably when Crawford applied these equations to a second sample and then combined the two samples a decrease in the amount of variance accounted for occurred resulting in 63% of the FSIQ variance accounted for by the combined NART-demographic equation, similar to the current study using the NART error only in the prediction equation. Crawford indicated that shrinkage in variance accounted for often occurs when combining samples but it is preferable to generate regression equations from combined standardisation and cross-validation samples as long as excessive shrinkage in variance does not occur.

Some studies in America have found the addition of demographic variables did not significantly increase the amount of variance accounted for in FSIQ (Blair & Speen, 1989; Bright, Jaldow & Koleman, 2002). It is likely that limited variability in education, SES and age specific to the samples studied may have reduced the predictive accuracy in these studies.

Willshire, Kinsella and Prior (1991) studied an Australian population of normal and clinical groups and found that for the total sample the inclusion of education with the NART provided a substantially higher estimate of premorbid cognitive functioning than NART alone. The combined equations of NART error score and education accounted for 46% of the variance in WAIS-R FSIQ compared to only 26% of the variance by NART error score alone. They also found that the combined NART-education equation accounted for more FSIQ variance in the over 55years group than in the less than 55years group for the combined sample.

Together these results suggest that the findings often reflect the sample studied and may not generalise to other samples. Notably, developing prediction equations specifically for TBI patients with data from a large TBI population study, using a within-participants design, assists in minimising the differences which may be found across TBI samples.

Appendix A5 provides a Table showing obtained NART errors and corresponding predicted NART error scores using the equation developed in this study. This can be used in clinical practice to convert NART scores obtained soon after TBI into a predicted NART score which will provide the clinician with a more valid estimate of premorbid IQ. The scores presented in this table appear to reflect the influence of education noted in the present thesis and provide further support for the suggestion that cognitive reserve may provide protection from NART sensitivity. For example there was minimal discrepancy between obtained and predicted NART errors for participants obtaining above average IQ scores (<14 NART errors) but, in contrast, participants obtaining low average IQ scores (>32 NART errors) showed increasingly larger discrepancies between predicted and obtained NART errors suggestive of NART sensitivity to TBI in participants with fewer years of education.

Accuracy rates for the NART prediction equation in the current study were examined and results showed that the NART equation accurately predicted (+/- 1 error point) the 12 month NART score in 17% of cases for the total sample and predicted the 12 month NART within +-

3 error points of 48% of the sample. When the sample was divided into IQ groups (≤ 100 or > 100 split) 19% of cases were accurately predicted and 50% of cases were predicted within +- 3 error points for the higher IQ group. Eleven percent of cases were accurately predicted in the lower IQ group and 46% of cases were predicted within +- 3 error points, suggesting the equation is slightly more accurate for people with an IQ above 100. The equation appeared more likely to overestimate the improvement in NART performance for people with an IQ ≤ 100 and underestimate the improvement in NART performance for people with an IQ of > 100, indicating there was a regression towards the mean as observed in many studies examining prediction equations.

To examine if accuracy rates could be improved an alternate method for predicting NART, and in turn estimating FSIQ, following TBI was assessed. The method examined in the current study used mean change scores between the initial and 12 month NART assessments according to education level and age group (see Table 9.11, page 174). This showed that individuals with less education experience more NART sensitivity following TBI than individuals with more education. People in the lower education group aged between 16 – 24 years of age displayed the most sensitivity, showing a mean decrease of 6.33 errors by 12 months post-injury. Accuracy rates (+-1 error point) varied from 15% for the group aged 25 – 40 years with less years of education to 24% accuracy for the group aged 41 – 60 years with more years of education. While group numbers were relatively small (ranging from 13 – 39) the majority of accuracy rates were equal to or above 19% for the mean change error method of predicting the recovered NART score. Prediction of the 12 month NART within +- 3 error points ranged from 41% for the less educated group aged 16 – 24 years to 69% for the higher education group aged

61 - 80 years of age. This suggests that for some cases the mean change method may be the better alternative for predicting NART and therefore premorbid IQ and Table 9.13 will assist clinicians in this decision. However the prediction equation provides a valid and useful alternative for predicting NART scores, accounting for 64% of the variance.

9.4.1 Limitations

There were a significant number of participants (20%) who actually worsened in NART performance over time (increase in NART errors) and there a much higher rate of worsening performance was observed for the group with an estimated IQ of ≤ 100 (27%). In contrast 18% of the group with estimated FSIQ of > 100 showed worse NART performance over time, suggesting more unexpected fluctuations in performance for those with lower IQ. Some studies have noted fluctuations in performance for TBI participants across assessments and Riley and Simmonds (2003) observed decline in performance for approximately 15% of their severe TBI sample. They suggested that reasons for decline may be subsequent head injuries, non-neurological factors such as anxiety, depression and fatigue on the day of assessment or inter-rater reliability. Watt and O'Carroll (1999) noted that higher depression ratings were associated with increased NART errors scores and suggested that while previous research argued that NART was not affected by depression this research had only examined clinically depressed patients. They concluded that the combination of TBI and raised levels of depression may contribute to increased NART errors. However, previous research examining NART performance in TBI has not reported details regarding worsening NART performance over time. Notably there are very few longitudinal studies examining this and it is also likely the samples are not comparable as the current study obtained participants from a large population study, compared with relatively small samples used in other research.

In summary, this study has provided two methods to assist clinicians ascertain a more valid estimate of premorbid IQ in TBI, given the degree of NART sensitivity observed in Study 1. The current study is the first to examine a mean change method and a prediction equation, developed with data from a large Australian TBI population study. The within-participants design examining TBI patients over a 12 month period strengthens the validity of these findings, although further validation of these methods in large TBI samples is warranted.

Chapter 10 will provide an overview of the results from the three studies conducted for the present thesis and discuss the clinical implications of these findings as well as suggestions for future research.

CHAPTER 10

General Discussion

The estimation of premorbid IQ in TBI is an integral component of neuropsychological practice. Obtained scores on cognitive tests are generally compared with some estimate of premorbid intellectual functioning. This comparison is required in such cases as dementia, where diagnosis requires clear and convincing evidence of cognitive decline. The forensic setting also requires clear evidence of cognitive decline and this is particularly relevant to TBI. Information regarding the amount of cognitive decline and impairment may be relatively easy in severe TBI but more difficult to document in mild TBI (Franzen et al., 1997). Insurance claims and rehabilitation are all reliant on efficient, valid and reliable estimates of premorbid IQ to provide an accurate picture of the amount of cognitive decline incurred following injury. This information guides rehabilitation program development and expectations for improvement in the future and is critical in regards to positive outcomes for individuals sustaining a TBI and their family.

Detailed studies of the validity and reliability of estimates of premorbid IQ are therefore an important area of research. Historically, many methods have been used for the estimation of premorbid IQ with clinical judgement alone being the earliest and objective methods developed since. These objective methods were reviewed fully in Chapter 3 and in summary they can be grouped into three main categories: First the best-performance method utilising WAIS 'hold' subtests; secondly, the measurement of current skills such as reading or WAIS Vocabulary also utilising the 'hold' test concept; and thirdly multiple regression techniques using demographic data.

The use of WAIS sub-tests utilising the 'hold' test concept (Lezak et al., 2004) has been criticised rather extensively for reliance on the score of one test only and the assumption this will provide a good estimate of premorbid IQ. These estimates do not account for the finding that there is extensive intra-individual sub-test scatter (Mortenson et al., 1991) or that there are low correlations between subtests, for example 0.41 between Object Assembly and Vocabulary (Wechsler. 1981). Several researchers have suggested that all WAIS subtests can be affected by brain injury to some degree or another, suggesting this method does not provide a valid estimate of premorbid IQ (Russell, 1972; Vanderploeg and Schinka, 1995).

Regression equations based on demographic variables have also fallen short as estimates of premorbid IQ. Studies have shown these equations account for varied amounts of WAIS FSIQ depending on the sample studied and range from 36% - 53% (reviewed in Chapter 3) Criticisms of this method include regression towards the population mean, restricted IQ range and large standard error of measurement and there are questions regarding the predictive power of this procedure with inconsistent findings observed in the literature (Langeluddecke & Lucas, 2004; Vanderploeg, Schinka, & Axelrod, 1996). Combining demographic variables and current measures into regression equations has been shown to increase the amount of WAIS FSIQ variance accounted for up to 73% (Crawford et al, 1989d) but these methods experience the limitations of both the demographic measures and 'hold' tests combined.

Reading, also referred to as a 'hold' test, was thought to be a highly practiced and over-learned skill and once established can be maintained despite deteriorations in other areas of intellectual functioning (Nelson & McKenna, 1975). The development of reading tests as estimates of premorbid IQ were based on four premises: reading is highly correlated with intelligence in the general population; reading ability is more resistant to dementia than is the WAIS Vocabulary subtest; the reading of irregular, rather than regular, words is more resistant to cognitive decline; and word reading taps previous knowledge while minimizing the demands on current cognitive capacity (Franzen et al., 1997; Willshire et al., 1991).

Based on this information Nelson (Nelson, 1982; Nelson & Willison, 1991) developed the NART to assist clinicians in estimating premorbid IQ in patients with dementia. Early research found that NART fulfilled all four criteria listed above. NART was highly correlated with intelligence in the general population, loading highly on '*g*', $r = 0.85$ (Crawford et al., 1989c) and research findings have indicated NART was a better estimate of premorbid IQ than WAIS Vocabulary in dementia, suggesting it was more resistant to cerebral dysfunction (Crawford et al., 1988a). Finally the use of irregular words in the NART eliminates the need for application of grapheme-phoneme rules required for 'regular' words suggesting the reading of irregular words provides a more sensitive measure of previous familiarity with words rather than a measure of continuing ability to analyse a complex visual stimulus as required for 'regular' words (Nelson, 1982; Nelson & O'Connell, 1978).

Subsequently, the NART became one of the most commonly used estimates of premorbid IQ in neuropsychological practice and was referred to as the 'gold standard' (Crawford, Deary, Starr

& Whalley, 2001a). However, studies confirming the validity of the NART, particularly with clinical groups such as TBI, consisted of small sample sizes, mixed diagnoses/severity groups and small age ranges and recent studies have questioned the validity of NART as an estimate premorbid IQ and suggest it is sensitive to TBI and therefore can underestimate premorbid IQ (see reviews in Chapters 5 & 6). These studies have also been limited by small samples, mixed severity samples and ill-defined assessment time post injury. Few studies have examined NART sensitivity to TBI prospectively over a 12 month time frame with a repeated measures design and the present thesis has extended the literature by implementing this design and addressing the limitations noted using a large sample from an Australian population study.

Therefore of the focus of the present thesis was to examine the sensitivity of the NART and AUSNART to TBI and to assess their validity as estimates of premorbid IQ in an Australian TBI sample. Additionally, the thesis provided clinicians with methods which will assist in making more accurate estimates of premorbid IQ.

10.1 Overview of the findings

Study 1 examined 194 participants with TBI and the results confirmed the hypothesis that NART was sensitive to TBI if assessed soon after injury. NART sensitivity was influenced by age, education, occupation and to a lesser degree severity. It is important to note that while the number of errors for severity groups was in the expected direction (increased errors with increased severity) the group with severe TBI only showed significantly more errors than the group with mild TBI at 6 months post injury and this was due to the differential recovery pattern for severity groups. That is, all groups regardless of severity showed similar NART sensitivity to TBI within one month of injury but the pattern of recovery was slower for the severe group. Whereas the mild and moderate groups improved significantly by 6 months the severe group did not show any change in NART performance until 12 months and therefore the gap widened between the severe and mild group error scores at the 6 month time point. This study is the first prospective longitudinal study examining all three severity groups and provides evidence for NART sensitivity to TBI not only for severe injuries as shown by Riley and Simmonds (2003) but also for mild and moderate injuries.

Previous research findings regarding severity effects have been inconsistent and based on correlational designs, Watt and O'Carroll (1998) did not find an association between a measure of severity, the Glascow Coma Scale (GCS) and NART but Morris et al. (2005) did (see Chapter 2.4). Morris et al. did not find an association between the measure of severity used in the present thesis, post traumatic amnesia (PTA) and NART performance. The results of the current study showed an association between PTA and NART performance at 6 months post injury only. Ponsford et al. (1995) noted that PTA is generally considered a more sensitive measure of severity than GCS hence the reason for its use as a measure of severity in the current thesis. Also, GCS score are often not available, particularly for mild TBI and information regarding the first GCS and the time obtained is often unreliable. Inconsistent results in earlier studies were due to small mixed samples and the Morris et al. study reported an extremely long mean time between injury and assessment (7.1 years) suggesting many of the participants would have recovered from the TBI by the time of assessment and Watt and O'Carroll only assessed a small TBI group ($n = 25$) on one occasion only, at least 9 months post injury.

An alternative explanation for the results of the present thesis is that severity is related to NART performance when recovery from TBI has occurred but it is not when the NART is administered soon after injury. In the current thesis PTA was correlated with the 6 month NART but not the initial and 12 month. The longitudinal design implemented in this thesis provided a more thorough examination of the relationship between NART performance and severity of injury and showed different recovery patterns of NART performance according to severity group but other factors such as education and age appear more influential on NART performance.

Contrary to the hypothesis age was found to have a significant effect on NART performance and recovery. Younger people aged between 16 and 24 years showed poorer NART performance at all assessments than older people aged 61 – 80 years and showed more NART sensitivity to TBI, with significant improvement by 6 months observed. In contrast, people aged 61 – 80 years produced fewer NART errors than younger people and did not show the same degree of sensitivity with a much smaller improvement in NART performance over time. While this has not been observed in previous research, correlational analyses has been the only method of examining the relationship between age and NART performance suggesting the current study, using analysis of variance, provided a clearer picture of the relationship between age and NART. A possible explanation for the affect of age on NART performance and sensitivity to TBI observed in this study is that older people have experienced more exposure to these words and they are therefore more familiar with the words. Also it could be speculated that older people may have received a more traditional English education of better quality, regardless of years spent at school. Johnson et al. (2006) referred to education quality in contrast to education quantity (years of education) and suggested reading ability was a measure of educational quality.

As hypothesised education had a significant effect on NART performance at all assessments and the longitudinal design of the current research provided strong evidence for differential NART recovery paths according to years of education. People with ≤ 11 years education experienced a far greater degree of initial NART sensitivity to TBI than people with > 11 years education and therefore showed greater improvement in NART performance over time. A decrease of 3 errors by 6 months post injury was observed by the group with fewer years of education compared to a decrease of 1 error point for those with more years of education. This shows that people with fewer years of education show NART sensitivity to TBI if administered soon after injury resulting in an underestimate of premorbid IQ. Again previous research has shown an association between education and NART performance but not by using a longitudinal repeated measures design using ANOVAs as performed in the present thesis.

Occupation based SES also had a significant effect on initial NART performance and recovery showing a very similar pattern to education and given the known relationship between these variables this was to be expected. People in professional occupations produced significantly fewer NART errors than people working in un-skilled labour and they did not show a significant change in performance over time suggesting they did not exhibit NART sensitivity to TBI. In contrast the un-skilled labourers showed an improvement of 6 error points at 6 months post injury reflecting a high degree of NART sensitivity to TBI when assessed soon after injury, resulting in an underestimate of premorbid IQ.

The observed lack of NART sensitivity to TBI for those with more education and higher SES may be explained by research regarding 'cognitive reserve' (e.g., Stern, 2002; Staff et al., 2003; Andel et al., 2006). Stern noted that the concept of reserve had stemmed from observations that there did not appear to be a direct relationship between the degree of brain pathology or brain damage and the clinical manifestations of the damage observed. Staff et al. explained cognitive reserve as the apparent protection from the onset of cerebral disease and/or cognitive decline in old age. While the concept of reserve has been studied extensively in relation to Alzheimer's disease Stern stated it should be relevant to any situation where brain injury occurs. Education and SES have been found to contribute independently to increased reserve and as such are considered measures of cognitive reserve (Stern, 2002). Research on cognitive reserve provides a good explanation for some of the findings in the present study. Higher education may be protecting the individual from the cognitive effects of TBI which result in poorer NART performance when assessed soon after injury for those with fewer years education.

As hypothesised, gender did not influence performance on the NART, confirming previous findings of no gender differences for NART performance (Crawford, 1988b; Wiens et al., 1993).

Study1 addressed the limitations of previous research and expanded on current knowledge by examining a large sample and individual severity groups (mild, moderate and severe) over a 12-month period, with specified assessment times post injury. The results highlight the need for clinicians to be aware of situations when NART will underestimate premorbid IQ as this will influence decisions regarding the degree of impairment for an individual with a TBI and the development of a rehabilitation program.

Given the concern noted regarding the validity of the NART as an estimate of premorbid IQ with Australian participants, Study 2 examined the AUSNART under the same conditions as outlined for Study 1. Also a comparison of NART and AUSNART performance was undertaken in a sub-sample of participants who completed both test at all three assessments. Overall the AUSNART performed similarly to NART and the results validated its characteristics compared to NART with high correlations, r =.80 - .92, observed between the tests. Both tests showed initial sensitivity to TBI and both were influenced significantly by education.

AUSNART was found to produce significantly higher estimates of premorbid IQ than the NART as hypothesised. Given the AUSNART equations were developed on a high education group it is likely to have inflated AUSNART estimates of premorbid IQ, although the sample studied in the current thesis also consisted of a slightly higher percentage of participants with more years of education which would also result in a mean estimated IQ higher than the population mean. AUSNART did show similar NART sensitivity to TBI when administered soon after injury, although recovery appeared slower, with a decrease in errors scores not observed until 12 months post injury, compared to 6 months for the NART.

In contrast to Study 1 and the hypotheses, injury severity did not influence AUSNART performance or recovery and the effect of age only showed a trend towards significance. In regards to age an explanation for the results may be that there were smaller age group numbers resulting in less power to detect a result. Alternatively, it is also possible that the AUSNART words are more familiar to younger Australian participants than the NART words and more culturally relevant, as the AUSNART words were chosen more recently and developed with a young Australian population. In comparison the NART word list developed, much earlier was not as familiar and resulted in poorer performance for young people.

As hypothesised and observed in Study 1, education had a significant effect on AUSNART performance and recovery over time. The results of Study 2 showed that people with more years of education showed no significant change in NART performance over 12 months. This provides further support for the suggestion that higher education may provide cognitive reserve which provides protection from some of the effects of brain injury, resulting in limited or no AUSNART sensitivity to TBI.

The hypothesis that gender would not influence AUSNART performance was supported in this study. Gender differences have not been examined for the AUSNART in previous research and this study provides the first examination of gender effects for AUSNART.

In summary, the AUSNART may be a more appropriate estimate of premorbid IQ in an Australian TBI population as it is a more culturally relevant reading test for all ages. However, further research is required as the present study provides the first available information regarding AUSNART performance in an Australian TBI population and the influence of demographic variables for this population. Further normative studies and studies with clinical samples are required to validate the AUSNART as an estimate of premorbid IQ in Australian TBI samples and to assist interpretation of the results of the present thesis.

As NART is currently accepted as the 'gold standard' in neuropsychological practice for the estimation of premorbid IQ methods which may assist clinicians to predict a more accurate NART score following TBI were examined in Study 3. The prediction equation using the initial 'impaired' NART as a predictor of the 12 month 'recovered' NART score accounted for 64% of the variance in the recovered NART but contrary to expectations the addition of demographic variables, such as education and age did not increase the amount of variance accounted for. Similar results have been found in other studies (Blair & Spreen, 1989, Bright et al., 2002) and it is likely that the amount of variance accounted for by education and age was explained in the initial NART score. Crawford et al. (1989d) found no additive effect of education and suggested it was due to the amount of covariance between NART and education, providing further support for the findings of this thesis. Also, it appears that inconsistent results in regards to the amount of NART and WAIS FSIQ variance accounted for by demographic variables are specific to the samples studied, with varying age ranges and years of education included. Therefore, while the NART prediction equation in the present thesis was developed with participants from a large TBI population study, further replication studies with large TBI samples is warranted.

Examination of accuracy rates found the equation developed in this study accurately predicted (within +- 1 error point) the improved 'unimpaired' NART in 19% of cases and within + - 3 errors for 48% of the cases. In 50% of cases the 12 month NART was predicted within +- 3 error points for the higher IQ group (estimated FSIQ >100) and 46% of cases were predicted within +- 3 error points for the low IQ group (estimated FSIQ ≤100), suggesting the equation is slightly more accurate for people with an IQ above 100. Notably as observed for FSIQ prediction equations in general the equation was more likely to overestimate NART improvement in people with lower IQ and underestimate NART improvement in people with higher IQ (see Chapter 3)

A mean change score method of predicting NART sensitivity to TBI according to education and age was also examined and showed that people with less than or equal to 11 years education aged between 16 – 24 years experienced the greatest degree of NART sensitivity to TBI showing a mean change score of 6 error points by 12 months post-injury. People with more than 11 years education and aged between 61 – 80 years experienced the least NART sensitivity, producing a mean change of 2 error points over 12 months. Accuracy rates (+- 1 error point) for predicting the 12 month NART ranged from 15% to 24 % with the majority of education/age groups producing accuracy rates of 19% and predictions within +- 3 error points ranged from 41% to 69% above suggesting this is a valid alternative to the prediction equation for estimating NART performance in TBI. However, clinicians will be required to judge which method is more accurate for the individual case, taking into account the person's years of education and age.

.

When estimating premorbid IQ in TBI, Table 9.13 (p. 176) and the Predicting NART error score Table provided in Appendix A5 will guide clinicians in making more accurate estimates of premorbid IQ in TBI. Education and age significantly influence NART performance following TBI and the two prediction methods provided assist in accounting for these factors when using NART to estimate premorbid IQ. While severity of injury did not show a strong effect on the initial 'impaired' or the 12 month 'recovered' NART, differential recovery paths were observed for severity groups according to education. Participants with fewer years of education showed more initial NART sensitivity to TBI than those with more years of education. However those with more education showed similar initial NART performance regardless of severity but those with a severe injury did not show the same improvement in errors as the mild and moderate group showed at 12 months post injury, suggesting severity of injury needs to be considered when using the prediction methods proposed in the current thesis.

An unexpected finding in the present thesis was that some participants actually showed poorer NART performance over time. Varied performance over time and worsening NART performance was also observed by Riley and Simmonds (2003), with 15% of their severe TBI sample showing worsening performance over time. They suggest this may be due to factors such as subsequent head injuries, which are common in this population but often not reported, and depression, fatigue and secondary ongoing health issues, which may influence performance on cognitive tasks. This provides further support for validating the equation developed in this study on other large TBI samples.

The present thesis has confirmed previous research suggestions that NART is sensitive to TBI (Freeman & Godfrey, 2000; Freeman, Godfrey, Harris and Partridge, 2001; Riley & Simmonds, 2003; Skilbeck, Allen & Brechin, 2005). It has expanded the knowledge in the field by providing more detailed information utilising a large longitudinal study of TBI patients, examining severity groups and controlling for assessment time post injury. In addition the present thesis provided a thorough examination of AUSNART with an Australian TBI population and is the first detailed study regarding its use with TBI participants. Further to this, the current thesis provided methods to assist clinicians make more accurate estimates of premorbid IQ, an integral component of neuropsychological practice.

10.2 Clinical implications

The estimation of premorbid IQ is a necessary and vital component of neuropsychological assessment in clinical practice. In the clinical setting it is paramount that the validity of an estimate of premorbid IQ is demonstrated in clinical samples such as TBI. The finding that NART and AUSNART performance can be sensitive to TBI and underestimate premorbid IQ, particularly if the person is aged between 16 – 24 year and has less than 12 years of education, has important implications for clinical practice. Underestimating premorbid IQ leads to an underestimate of cognitive decline following brain injury. Establishing premorbid IQ provides a base from which to interpret the performance, impaired or not impaired, of an individual on a range of cognitive tests. Research examining patients with TBI has consistently found impaired performance on tests of attention, memory, information processing, working memory and executive functioning (Carroll et al., 2004; Chan, 2000; Ponsford et al., 1995). To adequately assess if a person's performance is impaired on tests assessing these cognitive domains a valid estimate of premorbid IQ is required to provide a comparison.

The estimation of premorbid functioning should not rely on one method alone as many factors will influence a person's potential cognitive ability. Factors such as family breakdown, trauma, and psychological issues such as anxiety, depression and schizophrenia will impact on an individual's ability to perform to their potential. Basing estimation on a person's years of education or occupation attained alone without considering other factors may lead to misleading conclusions, underestimating premorbid functioning and therefore underestimating the gains which could be made in rehabilitation. Objective measures such as the NART assist the clinician estimate premorbid IQ in TBI and a thorough knowledge of its sensitivity to TBI as demonstrated in the present thesis has extended the available literature. Along with clinical judgement, objective measures remain an integral part of the estimation of premorbid IQ.

10.3 Future Directions

Further NART TBI studies, utilising a longitudinal repeated measures design are required to validate the current findings. The present thesis extended previous research by studying a large sample, examining and comparing three severity groups and assessing the same participants on three occasions over a 12 month period. Previous research has generally consisted of smaller samples, only examined one severity group or had small mixed severity group and have not utilised the prospective longitudinal design. However replication of this study with another large TBI population study would provide further validation of the findings.

There is a need for a large normative study of the AUSNART which will assist interpretation of the results of Study 2 of the present thesis. The results of Study 2 provide an examination of AUSNART performance in TBI and the longitudinal within participants design provides an excellent method to compare NART and AUSNART performance. However, it will be important to a gain a better understanding of AUSNART performance in healthy participants and to investigate further whether the equation developed by Hennessey and Mackenzie (1995) and used in the present thesis can be validated.

The current study involved approximately 15 trained research assistants who scored the NART and AUSNART responses. This adds to the robustness of the findings and while inter-rater reliability issues are not thought to be of significance for the NART it is possible. However, to ensure inter-rater reliability future research may benefit from taping the participants' NART and AUSNART responses to be scored later by a 'blind' rater. Multiple raters could be used and an inter-rater reliability study could be performed, reviewing the taped responses and comparing individual raters scoring.

Additionally, as noted by Wilshire et al. (1991) using the pronunciation guide in the most recent version of the Macquarie Dictionary provides a more appropriate assessment of the pronunciation of NART words in the Australian context as it provides alternative pronunciations commonly used in Australia making it more culturally relevant. This should be considered when conducting research with an Australian population.

In conclusion the prediction methods provided in the current thesis require cross-validation with other large TBI populations. However, given the results of this thesis, showing NART is sensitive to TBI and can therefore underestimate premorbid IQ, the NART prediction equation and mean change score method provided will assist clinicians to make more accurate estimates of premorbid IQ; an integral component in the development of rehabilitation programs and hence expectations for future positive outcomes for patients with TBI.

References

Alcott, D., Swann, R., & Grafham, A. (1999). Brief report: The effect of training on inter-rater reliability on the scoring of the NART. *British Journal of Clinical Psychology, 38,* 431-435.

Ander, R., Vigen, C., Mack, W.J., Clark, L.J. & Gatz, M. (2006). The effect of education and occupational complexity on rate of cognitive decline in Alzheimer's patients. *Journal of the International Neuropsychological Society, 12*, 147–152.

Ahmed, S., Bierley, R., Sheikh, J.I & Date, E.S. (2000). Post-traumatic amnesia after closed head injury: a review of the literature and some suggestions for further research. *Brain Injury*, 14, 765-780.

Australian Standard Classification of Occupations (ASCO), second edition (1997). Retrieved January 10, 2005, from
http://www.abs.gov.au/ausstats/abs@.nsf/0/A630E3FD6A69F7CCCA25697E001851DA

Baddeley, A., Emslie, H., & Nimmo-Smith, I. (1993). The Spot-the-Word test: A robust estimate of verbal intelligence based on lexical decision. *British Journal of Clinical Psychology,32,* 55-65.

Barona, A., Reynolds, C.R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. *Journal of Consulting and Clinical Psychology, 52,* 885-887.

Beardsall, L. (1998). Development of the Cambridge Contextual Reading Test for improving the estimation of premorbid verbal intelligence in older persons with dementia. *The British Journal of Clinical Psychology, 37*, 229-240.

Beardsall, L., & Brayne, C. (1990). Estimation of verbal intelligence in an elderly community: a prediction analysis using a shortened NART. *British Journal of Clinical Psychology, 29,* 83-90.

Beardsall, L., & Huppert, F. (1997). Short NART, CCRT and Spot-the-Word: comparisons in older and demented persons. *British Journal of Clinical Psychology, 36,* 619-622.

Berry, D.T.R., Carpenter, G.S., Campbell, D.A., Schmidt, F.A., Helton, K. & Lipke-Molby, T. (1994). The New Adult Reading Test-Revised: accuracy in estimating WAIS-R IQ scores obtained 3.5 years earlier from normal older persons. *Archives of Clinical Neuropsychology, 9*, 239-250.

Blair, J.R. & Spreen, O. (1989). Predicting premorbid IQ: a revision of the National Adult Reading Test. *The Clinical Neuropsychologist, 3*, 129-136.

Brayne, C., & Beardsall, L. (1990). Estimation of verbal intelligence in an elderly community: an epidemiological study using NART. *British Journal of Clinical Psychology, 29,* 217-223.

Bright, P., Jaldow, E., & Kopelman, M.D. (2002). The National Adult Reading Test as a measure of premorbid intelligence: a comparison with estimates derived from demographic variables. *Journal of the International Neuropsychological Society, 8*, 847-854.

Carrol, J.L., Cassidy, D., Holm, L., Kraus, J. & Conronado, V.G. (2004). Methodological issues and research recommendations for Mild Traumatic Brain Injury: the WHO collaborating task force on mild traumatic brain injury. Journal of Rehabilitation Medicine, *43*, 113 – 125.

Carstairs, J.R., Shores, A. (2000). The Macquarie Neuropsychological Normative Study (MUNNS): rationale and methodology. *Australian Psychologist, 35*, 36-40.

Cockburn, J., Keene, J., Hope, T., & Smith, P. (2000). Progressive decline in NART score with increasing dementia severity. Journal of Clinical and Experimental Neuropsychology, 22 (4), 508-517.

Crawford, J.R. (1992a). Current and premorbid intelligence measures in neuropsychological assessment. In Crawford, J.R., Parker, D.M., & McKinlay, W.W. (editors). *A handbook of neuropsychological assessment, pp* 21-49.

Crawford J.R. (1997). Provisional norms for NART estimated WAIS-III FSIQ. Personal communication.

Crawford, J.R., & Allan, K.M. (1997). Estimating premorbid WAIS-R IQ with demographic variables: regression equations derived from a UK sample. *The Clinical Neuropsychologist, 11*, 192-197.

Crawford, J.R., Allan, K.M., Cochrane, R.H.B., & Parker, D.M. (1990a). Assessing the validity of NART-estimated premorbid IQ's in the individual case. *British Journal of Clinical Psychology, 29,* 435-436.

Crawford, J.R., Besson, J.A.O., Parker, D.M., Sutherland, K.M., & Keen, P.L. (1987). Estimation of premorbid intellectual status in depression. *British Journal of Clinical Psychology, 26,* 313-314.

Crawford, J.R., Cochrane, R.H.B., Besson, J.A.O., Parker, D.M., & Stewart, L.E. (1990d). Premorbid IQ estimates obtained by combining the NART and demographic variables: construct validity. *Personality and Individual Differences, 11*, 209-210.

Crawford, J.R., Deary, I.J., Starr, J., & Whalley, L.J. (2001a). The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychological Medicine, 31*, 451-458.

Crawford, J.R., Hart, S., & Nelson, H.E. (1990b). Improved detection of cognitive reserve of cognitive impairment with the NART: an investigation employing hierarchical discriminant function analysis. *British Journal of Clinical Psychology, 29*, 239-241.

Crawford, J.R., Millar, J., & Milne, A.B. (2001b). Estimating premorbid IQ from demographic variables: a comparison of a regression equation vs clinical judgment. *British Journal of Clinical Psychology, 40 (1)*, 97-105.

Crawford, J.R., Nelson, H.E., Blackmore, L., Cochrane, R.H.B., & Allan, K.M. (1990c). Estimating premorbid intelligence by combining the NART and demographic variables: an examination of the NART standardization sample and supplementary equations. *Personality and Individual Differences, 11*, 1153-1157.

Crawford, J.R., Parker, D.M., Allan, K.M., Jack, A.M., & Morrison, F.M. (1991). The short NART: Cross-validation, relationship to IQ and some practical considerations. *British Journal of Clinical Psychology, 30*, 223-229.

Crawford, J.R., Parker, D.M., & Besson, J.A.O. (1988a). Estimation of premorbid intelligence in organic conditions. *British Journal of Psychiatry, 153*, 178-181.

Crawford, J.R., Parker, D.M., Stewart, L.E., Besson, J.A.O., & De Lacey, G. (1989a). Prediction of WAIS IQ with the national Adult Reading Test: cross-validation and extension. *British Journal of Clinical Psychology, 28*, 267-273.

Crawford, J.R., Stewart, L.E., Cochrane, H.B., Foulds, J.A., Besson, J.A.O., & Parker, D.M. (1989b). Estimating premorbid IQ from demographic variables: regression equations derived from a UK sample. *British Journal of Clinical Psychology, 28,* 275-278.

Crawford, J.R., Stewart, L.E., Cochrane, H.B., Parker, D.M., & Besson, J.A.O. (1989c). Construct validity of the National Adult Reading Test: a factor analytic study. *Personality and Individual Differences, 10*, 585-587.

Crawford, J.R., Stewart, L.E., Garthwaite, P.H., Parker, D.M., & Besson, J.A.O. (1988b). The relationship between demographic variables and NART performance in normal participants. *British Journal of Clinical Psychology, 27,* 181-182.

Crawford, J.R., Stewart, L.E., Parker, D.M., Besson, J.A.O., & Cochrane, R.H.B. (1989d) Estimation of premorbid intelligence: combining psychometric approaches improves predictive accuracy. *Personality and Individual Differences, 10,* 793-796.

Delbridge, A. (1985). *The Macquarie dictionary (rev. ed.)* Dee Why, NSW: Macquarie Library Pty. Ltd.

Folstein, M., Folstein, S.E., McHugh, P.R. (1975). "Mini-Mental State" a Practical Method for Grading the Cognitive State of Patients for the Clinician. *Journal of Psychiatric Research, 12*, 189-198.

Fortune, N. & Wen, X. (1999). *The definition, incidence and prevalence of acquired brain injury in Australia.* AIHW cat. no. DIS 15. Canberra: AIHW.

Franzen, E.J., Burgess, E.J., & Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology, 12*, 711-738.

Freeman, J., & Godfrey, H. (2000). The validity of the NART-RSPM index in detecting intellectual decline following traumatic brain injury: A controlled study. *British Journal of Clinical Psychology, 39, 95 -103.*

Freeman, J., Godfrey, H.P.D., Harris, J.K.J., & Partridge, F.M. (2001). Utility of a demographic equation in detecting impaired NART performance after TBI. *British Journal of Clinical Psychology, 40,* 221-224.

Fromm, D., Holland, A.L., Nebes, R.D., & Oakley, M.A. (1991). A longitudinal study of word-reading ability in Alzheimer's disease: evidence from the National Adult Reading Test. *Cortex, 27,* 367-376.

Green, R.E.A., Melo, B., Christensen, B., Ngo, L. Monette, G. & Bradbury, C. (2008). Measuring premorbid IQ in traumatic brain injury: an examination of the validity of the Wechsler test of Adult Reading (WTAR). *Journal of Clinical and Experimental Neuropsychology, 30,* 163-172.

Griffin, S.L., Mindt, M.R., Rankin, E.J., Ritchie, A.J., & Scott, J.G. (2002). Estimating premorbid intelligence: Comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology, 17,* 497-507.

Grober, E., & Silwinski, M. (1991). Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology, 13,* 933-949.

Hennessey, M., & Mackenzie, B. (1994). AUSNART: The development of an Australian version of the NART. Australian Society of the Study of Brain Impairment: treatment and long-term outcomes. Proceedings of the 18th Annual Brain Impairment Conference, Hobart, Australia (Fourez, J. & Page, N. Eds).

Holdnack, J.A., Lissner, D., Bowden, S.C., & McCarthy, K.A.L. (2004). Utilising the WAIS-III/WMS-III in a clinical practice: update of research and issues relevant to Australian normative research. *Australian Psychologist, 39*, 220-227.

Johnstone, B., & Wilhelm, K.L. (1996). The longitudinal stability of the WRAT-R reading subtest: Is it an appropriate estimate of premorbid intelligence? *Journal of the International Neuropsychological Society, 2,* 282-285.

Khan, F., Baguley, I.J., & Cameron, I.D, (2003). Rehabilitation after traumatic brain injury. *Medical Journal of Australia, 178*, 290-295.

Klesges, R.C., Wilkening, G.N., & Golden, C.J. (1981). Premorbid indices of intelligence: a review. *Clinical Neuropsychology, 3*, 32-39.

Klesges, R.C., & Troster, A.J. (1987). A review premorbid indices of intellectual and neuropsychological functioning: what have we learned in the past five years? *The International Journal of Clinical Neuropsychology, 9*, 1-11.

Kolb, B. & Whishaw, I.Q. (2009). *Fundamentals of Human Neuropsychology* (6th ed.). New York: Worth Publishers.

Krull, K.R., Scott, J.G., & Sherer, M. (1995). Estimation of premorbid intelligence from combined performance and demographic variables. The Clinical Neuropsychologist, 9, 083-088.

Langeluddecke, P.M., & Lucas, S.K. (2004). Evaluation of two methods of estimating premorbid intelligence on the WAIS-III in a clinical sample. *The Clinical Neuropsychologist, 18,* 423-432.

Levin, H. S., O'Donnell, V.M., & Grossman, R.G. (1979). The Galveston Orientation and Amnesia Test: a practical scale to assess cognition after head injury. *The Journal of Nervous and Mental Disease,* 167, 675- 683.

Lezac, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Lucas, S.K., Carstairs, J.R., & Shores, A. (2003). A comparison of methods to estimate premorbid intelligence in an Australian sample: data for the Macquarie University Neuropsychological Normative Study (MUNNS). *Australian Psychologist, 38*, 227-237.

Mathias, J.L., Bowden, S.C., Bigler, E.D. & Rosenfeld, J.V. (2007). Is performance on the Wechsler Test of Adult Reading affected by traumatic brain injury? *British Journal of Clinical Psychology, 46*, 457-466.

Mathias, J.L., Bowden, S.C. & Barrett-Woodbridge, M. (2007). Accuracy of the Wechsler Test of Adult Reading (WTAR) and National Adult Reading Test (NART) when estimating IQ in a healthy Australian sample. *Australian Psychologist, 42*, 49-56.

McCrimmon, S., & Oddy, M. (2006). Return to work following moderate-to-severe traumatic brain injury. Brain Injury, 20, 1037-1046.

Mockler, D., Riordan, J., & Sharma, T. (1996). A comparison of the NART (restandardised) and the NART-R (revised). *British Journal of Clinical Psychology, 35,* 567-572.

Morris, P.G., Wilson, J.T., Dunn, L.T., & Teasdale, G.M. Premorbid intelligence and brain injury. *British Journal of Clinical Psychology, 44,* 209-214.

Mortensen, E.L., Gade, A., & Reinisch, J.M. (1991). A critical note on Lezak's 'best performance method' in clinical neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 13*, 361-371.

Nelson, H.E. (1982). The National Adult Reading Test (NART): Test manual. Windsor: NFER, 1 - 13.

Nelson, H.E., & McKenna, P. (1975). The use of current reading ability in the assessment of dementia. *British Journal of Social and Clinical Psychology, 14,* 259-267.

Nelson, H.E., & O'Connell, A. (1978). Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex, 14*, 234-244.

Nelson, H.E., & Willison, J.R. (1991). The National Adult Reading Test (NART): Test manual (2nd Ed). Windsor: NFER, 14 - 21.

O'Carrol, R.E., & Gilleard, C.J. (1986). Estimation of premorbid intelligence in dementia. *British Journal of Clinical Psychology, 25,* 1157-158.

O'Carrol, R.E. (1987). The inter-rater reliability of the National Adult reading Test (NART): A pilot study. *British Journal of Clinical Psychology, 26*, 229-230.

O'Carrol, R.E. (1995). The assessment of premorbid ability: a critical review. *Neurocase, 1,* 83-89.

Olver, J.H., Ponsford, J.L., & Curran, C.A. (1996). Outcome following traumatic brain injury: a comparison between 2 and 5 years after injury. *Brain Injury, 10*, 841-848.

Paolo, A.M., Troster, A.I., Ryan, J.J., & Koller, W.C. (1997). Comparison of NART and Barona demographic equation premorbid estimates in Alzheimer's disease. *Journal of Clinical Psychology, 53*, 713-722.

Patterson, K., Graham, N., & Hodges, J.R. (1994). Reading in dementia of the Alzheimer type: a preserved ability? *Neuropsychology, 8,* 395-407.

Perez, S.A. Schlottman, R.S., Holloway, J.A., Ozolins, M.S. (1996). Measurement of premorbid intellectual ability following brain injury. *Archives of Clinical Neuropsychology, 11 (6),* 491-501.

Ponsford, J., Sloan, S., & Snow, P. (1995). *Traumatic brain injury: rehabilitation for everyday adaptive living.* Psychology Press Ltd: East Susgender, UK.

Ponsford, J., Willmott, C., Rothwell, A., Cameron, P., Kelly, A-M., Nelms, R., & Curran, C. (2002). Impact of early intervention on outcome following mild head injury in adults. *Journal of Neurology, Neurosurgery and Psychiatry, 73,* 330-332.

Powell, B.D., Brossart, D.F., & Reynolds, C.R. (2003). Evaluation of the accuracy of two regression-based methods for estimating premorbid IQ. *Archives of Clinical Neuropsychology, 18,* 277-292.

Richards, M., & Sacker, A. (2003). Lifetime antecedents of cognitive reserve. *Journal of Clinical and Experimental Neuropsychology, 25 (5),* 624-624.

Riley G.A., & Simmonds, L.V. (2003). How robust is performance on the National Adult Reading Test following traumatic brain injury? *British Journal of Clinical Psychology, 42,* 319-328.

Rohling, M.L., Meyers, J.E., & Mills, S.R. (2003). Neurosychological impairment following traumatic brain injury: A dose response analysis. *The Clinical Neuropsychologist, 17,* 289-302.

Ruff, R. (2005). Two decades of advances in understanding of mild traumatic brain injury. Journal of Head Trauma Rehabilitation, 20, 5-18.

Russell, E.W. (1972). WAIS factor analysis with brain-damaged participants using criterion measures. *Journal of Consulting and Clinical psychology, 39,* 133-139.

Russell, A.J., Munro, J., Jones, P.B., Hayward, P., Hemsley, D.R., & Murray, R.M. (2000). The National Adult reading test as a measure of premorbid IQ in schizophrenia. *British Journal of Clinical Psychology, 39,* 297-305.

Savage, R. (1995). An educator's manual: What educators need to know about students with TBI. Washington, DC: Brain Injury Association.

Schmand, B., Geerlings, M.J., Jonker. C., & Lindeboom, J. (1998). Reading ability as an estimator of premorbid intelligence: does it remain stable in emergent dementia? Journal of clinical and Experimental Neuropsychology, 20, 42-51.

Schoenberg, M.R., Scott, J.G., Duff., K., & Adams, R.L. (2002). Estimation of WAIS-III intelligence from combined performance and demographic variables: development of the OPIE-3. *The Clinical Neuropsychologist, 16*, 426-438.

Schonell, F.J. (1966). *The psychology and teaching of reading*. London: Oliver & Boyd, Ltd.

Schlosser, D., & Ivison, D. (1989). Assessing memory deterioration with the Wechsler Memory Scale, the National Adult Reading Test, and the Schonell Graded Word Reading Test. *Journal of Clinical and Experimental Neuropsychology, 11*, 785-792.

Sharpe, K., O'Carroll, R. (1991). Estimating premorbid intellectual level in dementia using the National Adult reading Test: A Canadian study. *British Journal of Clinical Psychology, 30,* 381-384.

Shores, A., & Carstairs, J.R. (2000). The Macquarie University Neuropsychological Normative Study (MUNNS): Australian norms for the WAIS-R and WMS-R. *Australian Psychologist, 35*, 41-59.

Skilbeck, C., Allen, L., & Brechin, D. (2005). NART prediction and impairment in neurological patients. *Neuropsychological Rehabilitation, 15,* 69-75.

Snow, W., Tierney, M., Zoritto, M., Fisher, R., & Reid, D. (1989). WAIS-R test-retest reliability in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology, 11,* 423-428.

Staff, R.T., Murray, A.D., Deary, I.J. & Whalley, L.J. (2004). What provides cerebral reserve? *Brain,* 127, 1191 – 1199.

Stebbins, G.T., Wilson, R.S., Gilley, D.W., Bernard, B.A., Fox, J.H. (1990). Use of the National Adult Reading Test to estimate premorbid IQ in dementia. *The Clinical Neuropsychologist, 4,* 18-24.

Taylor, K.I., Salmon, D.P., Rice, V.A., Bondi, M.W., Hill, L.R., Ernesto, C.R., & Butters, N. (1996). Longitudinal examination of American National Adult Reading Test (AMNART) performance in dementia of the Alzheimer type (DAT): validation and correction based on degree of cognitive decline. *Journal of Clinical and Experimental Neuropsychology, 18,* 883-891.

Teasdale, G.M. (1995). Head injury. *Journal of Neurology, Neurosurgery and Psychiatry, 58,* 526-539.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: a practical scale. *The Lancet,* 81-83.

Turner, G.W. (1984). *The Australian Pocket Oxford Dictionary.* Melbourne: Oxford University Press.

Van den Broek, & Bradshaw, C.M. (1994). Detection of acquired deficits in general intelligence using the National Adult Reading test and Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology, 33,* 509-515.

Vanderploeg, R.D., & Schinka, J.A. (1995). Predicting WAIS-R IQ premorbid ability: combining subtest performance and demographic variable predictors. *Archives of Clinical Neuropsychology, 10,* 225-239.

Vanderploeg, R.D., Schinka, J.A., & Axelrod, B.N. (1996). Estimation of WAIS-R Premorbid Intelligence: current ability and demographic data used in a best-performance fashion. *Psychological Assessment, 8,* 404-411.

Vieth, A.Z., Johnstone, B., & Dawson, B. (1996). Extent of intellectual, cognitive and academic decline in adolescent traumatic brain injury. *Brain Injury, 10 (6),* 465-470.

Watt, K.J. & O'Carroll, R.E. (1999). Evaluating methods for estimating premorbid intellectual ability in closed head injury. *Journal of Neurology, Neurosurgery and Psychiatry, 66,* 474-479.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4^{th} ed.). Baltimore: Williams and Wilkins.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence-Revised (WAIS-R).* New York: The Psychological Corporation.

Wechsler, D. (1997). Wechsler Adult Intelligence and Memory Scales (WAIS-III; WMS-III). New York: Psychological Corporation.

Wechsler, D. (2001). *Wechsler Test of Adult Reading (WTAR).* New York: The Psychological Corporation.

Williams, J.M., Gomes, F., Drudge, O.W., & Kessler, M. (1984). Predicting outcome from closed head injury by early assessment of trauma severity. *Journal of Neurosurgery, 61*, 581-585.

Willshire, D., Kinsella, G., & Prior, M. (1991). Estimating WAIS-R IQ from the National Adult Reading Test: A cross validation. *Journal of Clinical and Experimental Neuropsychology, 13 (2),* 204-216.

Wilson, R.S., Rosenbaum, G., Brown, G., Rourke, D., & Whitman, D. (1978). An index of premorbid intelligence. *Journal of Consulting and Clinical Psychology, 46*, 1554-1555.

APPENDIX A

A1	NART Word List	204
A2	NART errors and estimated WAIS-III FSIQ scores	205
A3	AUSNART Word List	206
A4	AUSNART errors and estimated FSIQ scores	207
A5	Conversion of obtained NART score to predicted NART score	209

APPENDIX A1
NART WORD LIST

CHORD	SUPERFLUOUS
ACHE	SIMILE
DEPOT	BANAL
AISLE	QUADRUPED
BOUQUET	CELLIST
PSALM	FACADE
CAPON	ZEALOT
DENY	DRACHM
NAUSEA	AEON
DEBT	PLACEBO
COURTEOUS	ABSTEMIOUS
RAREFY	DETENTE
EQUIVOCAL	IDYLL
NAIVE	PUERPERAL
CATACOMB	AVER
GAED	GAUCHE
THYME	TOPIARY
HEIR	LEVIATHAN
RADIX	BEATIFY
ASSIGNATE	PRELATE
HIATUS	SIDEREAL
SUBTLE	DEMESNE
PROCREATE	SYNCOPE
GIST	LABILE
GOUGE	CAMPANILE

APPENDIX A2
NART ERRORS AND ESTIMATED WAIS-III FSIQ
Provisional Norms (Crawford, 1997)

NART Errors	WAIS-III FSIQ	WAIS-III VIQ
0	123.2	122.2
1	122.2	121.1
2	121.1	120.0
3	120.1	119.0
4	119.1	117.9
5	118.1	116.8
6	117.0	115.7
7	116.0	114.7
8	115.0	113.6
9	113.9	112.5
10	112.9	111.4
11	111.9	110.4
12	110.9	109.3
13	109.8	108.2
14	108.8	107.1
15	107.8	106.0
16	106.7	105.0
17	105.7	103.9
18	104.7	102.8
19	103.6	101.7
20	102.6	100.7
21	101.6	99.6
22	100.6	98.5
23	99.5	97.4
24	98.5	96.4
25	97.5	95.3
26	96.4	94.2
27	95.4	93.1
28	94.4	92.0
29	93.4	91.0
30	92.3	89.9
31	91.3	88.8
32	90.3	87.7
33	89.2	86.7
34	88.2	85.6
35	87.2	84.5
36	86.2	83.4
37	85.1	82.4
38	84.1	81.3
39	83.1	80.2
40	82.0	79.1
41	81.0	78.0
42	80.0	77.0
43	79.0	75.9
44	77.9	74.8
45	76.9	73.7
46	75.9	72.7
47	74.8	71.6
48	73.8	70.5
49	72.8	69.4
50	71.8	68.4

APPENDIX A3
AUSNART WORD LIST

ECHIDNA	COLONEL	CONFIDANT
OPAQUE	PTERODACTYL	ECHELON
UNCOUTH	SCEPTRE	PSYCHE
BOUQUET	VOYEUR	VISCOUNT
MAUVE	LINGERIE	CORPS
PLATEAU	CHASM	GOSHAWK
POLTERGEIST	CAMARADARIE	POSTHUMOUS
MEDIOCRE	EUCHRE	TSETSE
ENCORE	ALBEIT	QUADRUPED
SERGEANT	FAÇADE	COMMUNIQUE
GAOLED	SUPERFLUOUS	SCOURGE
PSALM	AEON	CACHE
PLACEBO	HIATUS	GAUCHE
CRUSTACEAN	CADAVER	PUERPERAL
ETIQUETTE	LOCALE	PENCHANT
HEIR	NONCHALANT	AVER
PHEASANT	STRYCHNINE	BEATIFY
RENDEZVOUS	TRIENNIAL	COELCANTH
ACACIA	AWRY	CABTABILE
CONNOISSEUR	CELLIST	BACCARAT
QUAY	GENRE	
SURETY	LEVIATHAN	

APPENDIX A4

AUSNART errors and estimated FSIQ scores*

AUSNART Score	Estimated FSIQ
1	134
2	134
3	133
4	132
5	131
6	130
7	130
8	129
9	128
10	127
11	126
12	125
13	125
14	124
15	123
16	122
17	121
18	120
19	120
20	119
21	118
22	117
23	116
24	116
25	115
26	114
27	113
28	112
29	111
30	111
31	110
32	109
33	108
34	107
35	107
36	106
37	105
38	104
39	103
40	102
41	102

42	101
43	100
44	99
45	98
46	97
47	97
48	96
49	95
50	94
51	93
52	93
53	92
54	91
55	90
56	89
57	88
58	88
59	87
60	86
61	85
62	84
63	83
64	83

*Estimated FSIQ based on equation from Hennessy & Mackenzie (1995) study
Estimated FSIQ = 135.27- .822 (AUSNART error score)

APPENDIX A5

Conversion of obtained NART error score to predicted NART error score*

Obtained NART error score	Predicted NART error score
1	3
2	3
3	4
4	5
5	6
6	6
7	7
8	8
9	9
10	10
11	10
12	11
13	12
14	13
15	13
16	14
17	15
18	16
19	16
20	17
21	18
22	19
23	19
24	20
25	21
26	22
27	23
28	23
29	24
30	25
31	26
32	26
33	27
34	28
35	29
36	29
37	30
38	31

39	32
40	32
41	33
42	34
43	35
44	36
45	36
46	37
47	38
48	39
49	39
50	40

*Equation developed in the present thesis (p. 170)
Improved 12 month NART Error = 1.861 + (initial NART error score x .765)